Title
Essays on the Incentives for Innovation and Voluntary Knowledge Transfer

Author
Kuo, Dennis William

Publication Date
2016

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
Los Angeles

Essays on the Incentives for Innovation and Voluntary Knowledge Transfer

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Economics

by

Dennis William Kuo

2016

ABSTRACT OF THE DISSERTATION

Essays on the Incentives for Innovation and Voluntary Knowledge Transfer

by

Dennis William Kuo

Doctor of Philosophy in Economics

University of California, Los Angeles, 2016

Professor Hugo Andres Hopenhayn, Chair

In the following essays I study the determinants of firms' incentives to innovate and voluntarily transfer knowledge to other firms.

Technology licensing and inventor job transitions are two examples of knowledge diffusion that takes place voluntarily between firms in a market. In this context, the incidence of transfer will depend on product market competition. I ask how changes in intellectual property policies affect voluntary knowledge transfer and innovation across different degrees of product substitutability. I also investigate the empirical relationship between the incidence of knowledge transfer and substitutability and compare that relationship against existing theories of knowledge transfer.

In the first chapter, I use a two-stage duopoly game of innovation and knowledge transfer to show that innovator bargaining power determines the relationship of innovation to the substitutability of the competitors' products. In particular, innovation increases in substitutability when the innovator's bargaining power is low. In such a situation, the model predicts that the incidence of knowledge transfer will first rise and then fall as a function of substitutability. I establish these results using reduced form assumptions on profit, but also show that these assumptions are consistent with an environment of nested CES demand, price competition, and constant marginal cost-reducing innovation.

In the second chapter, I find that the predicted non-monotonic pattern of knowledge transfer

holds empirically between pairs of firms. Specifically, a rising-then-falling relationship exists in the incidence of both technology licensing deals and inventor job transitions as a function of firms' bilateral product market overlap. In a cross-industry sample of firms, I measure product market overlap between two firms as the correlation between the distribution of shares of their total sales across SIC 4-digit industries. I find the rising-then-falling relationship between knowledge transfer and market overlap in this cross-industry sample after controlling for bilateral technological overlap between firms' patent portfolios. This empirical finding isolates the strategic competitive determinants of knowledge transfer and shows that they are economically significant. The results also constitute indirect evidence for the existence of compensation mechanisms that internalize the knowledge spillovers from R&D worker job mobility.

In the third chapter, I find that an infinite-horizon dynamic duopoly game confirms the non-monotonic empirical pattern at low innovator bargaining power. In this dynamic game, a technology leader and follower can be separated by any number of innovations. I use the dynamic model to show that greater bargaining power positively impacts the output growth rate through increased innovation. However, raising the bargaining power also generates a countervailing shift away from neck and neck innovation; this shift has a negative impact on growth and the net result is ambiguous.

The dissertation of Dennis William Kuo is approved.

Nico Voigtlaender

Pablo David Fajgelbaum

John William Asker

Hugo Andres Hopenhayn, Committee Chair

University of California, Los Angeles

2016

To my parents and sister, who patiently listened, encouraged, and supported me every step of the way.

TABLE OF CONTENTS

List of Figures

List of Tables

ACKNOWLEDGMENTS

I am deeply indebted to Hugo Hopenhayn for his invaluable guidance, mentorship, and tireless encouragement. John Asker and Pablo Fajgelbaum provided essential advice for this project. I would also like to thank Andy Atkeson, Leah Boustan, Ariel Burstein, Fernando Giuliano, Federico Grinberg, Andreas Gulyas, Ryan Martin, Lee Ohanian, Mark Schankerman, Liyan Shi, Connan Snider, Nico Voigtlaender, Gabriel Zaourak, and seminar participants at UCLA for very useful comments and suggestions.

VITA

2005	B.S. (Electrical Engineering), Stanford University.
2007	M.S. (Management Science and Engineering), Stanford University.
2012	M.A. (Economics), University of California, Los Angeles.

CHAPTER 1

Innovation and Knowledge Transfer in a Duopoly Game

1.1 Introduction

Technology diffusion between firms contributes to productivity improvements - Shapiro (1985) terms this the ex post effect of diffusion, and empirical examples are found in Irwin and Klenow (1994) and Bloom et al. (2013). At the same time, if technology diffuses between two firms that compete in related markets, it will improve the product market position of the recipient at the expense of the original innovator. The incidence of knowledge transfer and its impact on innovation incentives - what Shapiro (1985) terms the ex ante effect of diffusion - will depend on product market competition and the degree to which spillovers are internalized. Two prominent channels of inter-firm knowledge diffusion serve as examples of voluntary transfers in a market for knowledge: technology licensing deals and inventor job transitions.[1] In this context of voluntary transfers, I ask how the innovator's ability to appropriate the gains from trade in knowledge affects the transfer of knowledge and the incentives for innovation. The effects of appropriability vary with the degree of product market substitutability between firms, and the model admits the interesting possibility that the incidence of knowledge transfer first rises and then falls as a function of product substitutability.

I develop a two-stage duopoly model of innovation and knowledge transfer. In the first stage of the model, two symmetric firms decide whether to innovate by paying a fixed cost. This innovation is a technology that can be used by both firms, even if they produce differentiated products.

If only one firm innovates in the first stage, and there are gains from trade in knowledge,

[1] An inventor's current employer can counterbid in the face of a poaching attempt.

then in the second stage the duopolists Nash bargain over the lump sum payment from imitator to innovator. Both firms end up with the innovation. The bargaining power of the innovator serves as an index of appropriability. The firms make profits in the product market that depend on their product substitutability as well as their final technological positions.

The analysis in this paper focuses on a symmetric, mixed-strategy equilibrium in which firms choose a probability of innovation. Such an equilibrium exists when the fixed innovation cost falls in an intermediate range. Analysis of this equilibrium yields a tractable intuition for the incentives faced by firms engaged in stochastic innovation.

When the products of these competing duopolists are highly substitutable, there will be no knowledge transfer because the innovator charges more for the innovation than the buyer is willing to pay. The difference between these two amounts is the gains from trade in knowledge, or the surplus. Under reasonable assumptions on profits, surplus will decrease from positive to negative as product substitutability increases. When products are highly substitutable and surplus is negative, there will be no knowledge transfer because the innovator possesses the means to exclude.

At lower degrees of product substitutability, the bargained apportionment of surplus will determine how the probabilities of innovation and knowledge transfer change as substitutability increases. Suppose the innovator possesses no bargaining power; the probability of innovation is low because a non-innovator captures the surplus from knowledge transfer and essentially free-rides. Holding fixed a low innovator bargaining power, the incentive to innovate increases with substitutability as the payoff from free-riding off surplus decreases in substitutability.[2]

When innovator bargaining power is low, the ex-ante probability of knowledge transfer can first rise and then fall as a function of product substitutability. As innovation increases in substitutability, so does the probability of an asymmetric innovation outcome, which coincides with knowledge transfer when surplus is positive.[3] Thus at a low innovator bargaining power, the incidence of knowledge transfer initially increases in substitutability until surplus becomes

[2]The gain to pulling ahead increases in substitutability, which reinforces this outcome.

[3]Innovation and knowledge transfer both increase in substitutability only when the probability of innovation starts well below $\frac{1}{2}$.

negative, at which point the incidence of transfer falls to zero.

In using technology licensing as a motivation and application, this paper connects to a voluminous literature on licensing. Katz and Shapiro (1985) focus on fixed-fee licensing, citing the difficulty in monitoring and enforcing per-unit royalties. They find that raising the licensor's bargaining power increases the incentives for innovation. However, the setting is one where an upstream research laboratory develops and sells an innovation to one of two competitors, which then decides whether to license to its rival. Katz and Shapiro (1987) find that low licensor bargaining power can induce a firm to wait for its rival to innovate, thus delaying the development of an innovation. My paper shares these qualitative results but focuses on how the role of bargaining power varies with competitors' product substitutability. In doing so, it permits a comparison to and explanation of the empirical relationship between knowledge transfer and product substitutability.

Arora and Fosfuri (2003) analyze a setting in which firms occupy niches with differentiated products that also correspond to differentiated technologies. Innovators can license to entrants, but these entrants will then necessarily occupy the same product-technology niche as the licensor. Empirically, their assumption of product-technology niches is supported by a paper such as Fosfuri (2006) that uses licensing data from the chemical industry. However, Bloom et al. (2013) show more broadly that there is meaningful independent variation in firms' product market overlap, apart from their degree of technological similarity. This paper contributes to the literature by evaluating bilateral product market substitutability as an independent factor in determining licensing and innovation, apart from technological variation.

Finally, Pakes and Nitzan (1983) lay the foundation for interpreting the labor mobility of R&D workers as bilateral sales of knowledge between firms. In particular, a backloaded compensation scheme exists that forces a scientist to internalize the knowledge spillover created by their departure to create a spin-out. This compensation scheme is isomorphic to a bilateral fixed-fee licensing deal between the parent firm and the spin-out. This correspondence between the compensation scheme and licensing allows the analysis in my paper to be used in predicting labor mobility patterns as a function of source and destination firm substitutability. Hamasaki (2012) and Fosfuri and Rønde (2004) consider backloaded compensation together with firm location decisions that are affected by non-compete and trade secret laws, respectively. They do not characterize endogenous innovation as a function of a continuous measure of market competition. Neither do they address how this relationship would be altered by changing the degree of innovator appropriability afforded by the backloaded compensation contract. On the other hand, Heggedal et al. (2014) do analyze how compensation affects firms' innovation incentives when poaching is possible, but they abstract from strategic product market considerations.

The rest of the paper is structured as follows. Section 2 sets out the two-stage duopoly game environment and characterizes knowledge transfer surplus. Section 3 presents the pure and mixed strategy equilibria that can occur in the game. Section 4 proceeds with the mixed strategy equilibrium and demonstrates that bargaining power changes the relationship of innovation to product substitutability. Section 5 shows that the incidence of knowledge transfer can be first rising and then falling in product substitutability. Section 6 proves that the stylized assumptions on profit made in the analysis are consistent with a nested CES demand system. Section 7 concludes.

1.2 Game Environment and Surplus

There are two stages in the game:

1. In the first stage, duopolists start at a normalized technology level of 0 and decide whether to take action $s = I$ (for Innovate) at cost c or action $s = N$ (for No Innovation).

2. In the second stage, if there are gains from trade in technology, then the innovator bargains with the imitator over the size of a lump sum payment. The innovator's bargaining weight is given by θ.

After the second stage of the game, duopolists compete in the product market with their final technology levels.

The timing is illustrated in Figure 1.1.

Duopolist i makes profit $\pi(H_i, H_{-i}; \sigma)$ where $H_i \in \{0,1\}$ is the technology level of firm i and H_{-i} is the rival's technology level. The degree of competition is given by the index $\sigma \in [0,1]$. Because the number of firms is fixed at two, σ specifically represents substitutability, where $\sigma = 1$ represents perfect substitutes and $\sigma = 0$ represents independent demand. I assume that $\pi(H_i, H_{-i}; \sigma)$ is continuous in σ.

The possible technology outcomes that can be realized in this game are given by:

$$\pi(H_i, H_{-i}) \in \{\pi(0,0), \pi(0,1), \pi(1,0), \pi(1,1)\}.$$

I assume that there are gains to relative and level increases in the technology level of firm i, which yields the following relations:

$$\pi(1,0) \geq \pi(1,1) \geq \pi(0,0) \geq \pi(0,1), \forall \sigma.$$

The innovation undertaken in the first stage is of potential use to both firms, even if they produce differentiated products. As such, gains from trade may exist in Stage 2 if firms take asymmetric actions $(s_i, s_{-i}) \in \{(I,N), (N,I)\}$ in Stage 1. Knowledge transfer is conducted as a sale of technology. The benefit to the imitator of acquiring the technology is given by $\pi(1,1) - \pi(0,1)$; this is also their willingness to pay. The damage done to the innovator in the product market is $\pi(1,0) - \pi(1,1)$ and represents their reservation value. The surplus is given by the difference between the imitator's willingness to pay and the innovator's reservation value:

$$S = (\pi(1,1) - \pi(0,1)) - (\pi(1,0) - \pi(1,1)).$$

Knowledge transfer occurs if and only if $S > 0$, and in this case the innovator and imitator bargain over the division of surplus. Bargaining determines the size of the lump sum payment made from imitator to innovator. Suppose the innovator has a bargaining weight $\theta = 0$ corresponding to minimal intellectual property protection. In this case, the innovator charges their reservation value $\pi(1,0) - \pi(1,1)$. The net change in profit for the innovator is $\theta S = 0$, and the net change for the imitator is $(1-\theta)S = S$. Alternatively, if the innovator has full bargaining weight $\theta = 1$, then the innovator charges the imitator's full willingness to pay of $\pi(1,1) - \pi(0,1)$. The net change in profit for the innovator is $\theta S = S$, and for the imitator it is $(1-\theta)S = 0$.

To understand how surplus behaves as substitutability σ increases, I decompose S into two bracketed terms as follows:

$$S = \{\pi(1,1) - \pi(0,0)\} + \{(\pi(1,1) - \pi(0,1)) - (\pi(1,0) - \pi(0,0))\}.$$

The first term $\{\pi(1,1) - \pi(0,0)\}$ in brackets is the rents to level innovation, which I assume decreases from positive to zero[4] in σ:

$$\begin{aligned}
\{\pi(1,1) - \pi(0,0)\}|_{\sigma=0} &> 0, \\
\{\pi(1,1) - \pi(0,0)\}|_{\sigma=1} &= 0, \\
\frac{d\{\pi(1,1) - \pi(0,0)\}}{d\sigma} &< 0.
\end{aligned}$$

I refer to the second term in brackets $\{(\pi(1,1) - \pi(0,1)) - (\pi(1,0) - \pi(0,0))\}$ as the strategic substitutes effect, because I assume that the effect of firm $-i$'s innovation on firm i's incentives to innovate decreases from zero to negative in σ:

$$\begin{aligned}
\{(\pi(1,1) - \pi(0,1)) - (\pi(1,0) - \pi(0,0))\}|_{\sigma=0} &= 0, \\
\{(\pi(1,1) - \pi(0,1)) - (\pi(1,0) - \pi(0,0))\}|_{\sigma=1} &< 0, \\
\frac{d\{(\pi(1,1) - \pi(0,1)) - (\pi(1,0) - \pi(0,0))\}}{d\sigma} &\leq 0.
\end{aligned}$$

[4]The zero-endpoint assumption is sufficient but not necessary for the proposition that follows, and corresponds to Bertrand competition and constant marginal cost.

In Section 1.6 I show that these assumptions on the ranking of the profitability of technology outcomes, the rents to level innovation, and the strategic substitutes effect are satisfied in an environment consisting of nested CES demand, price competition, and constant marginal cost-reducing innovation.

In Appendix Section 1.9.1 I prove the following proposition:

Proposition 1. *Given the assumptions above on the two terms that compose surplus, S decreases from positive to negative in σ:*

$$\begin{aligned} S|_{\sigma=0} &> 0, \\ S|_{\sigma=1} &< 0, \\ \frac{dS}{d\sigma} &< 0. \end{aligned}$$

1.3 Equilibria

The payoffs from the perspective of Stage 1 are given in Table 1.1 where $\bar{S} = \max\{0, S\}$.

The equilibrium is determined by the magnitude of the innovation cost c. If c is greater than the gain from playing I vs. N regardless of what the rival firm plays, then firms play (N, N). The cutoff $\bar{c}$ for this no-investment region is given by:

$$\bar{c} = \pi(1,0) + \theta\bar{S} - \pi(0,0).$$

This is equal to the gain from playing I when the rival plays N.

Likewise, if c is lower than the gain from playing I vs. N regardless of the rival's action, then firms play (I, I). The cutoff $\underline{c}$ for this region is given by:

$$\underline{c} = \pi(1,1) - \left(\pi(0,1) + (1-\theta)\bar{S}\right).$$

This is equal to the gain from playing I when the rival plays I.

When $c \in \left(\underline{c}, \bar{c}\right)$, asymmetric equilibrium strategies $(s_i, s_{-i}) \in \{(I,N),(N,I)\}$ are played. The symmetric equilibrium is a mixed-strategy probability p of playing I, which is given by:

$$p = \frac{\pi(1,0) - \pi(0,0) + \theta\bar{S} - c}{\pi(1,0) - \pi(0,0) - (\pi(1,1) - \pi(0,1)) + \bar{S}}.$$

Overall, for any given innovation cost c, a higher innovator bargaining weight θ raises the cutoffs $\underline{c}$ and $\bar{c}$ and increases the likelihood of innovation. The symmetric equilibria are depicted as a function of c in Figure 1.2.

In what follows, I assume that $c \in \left(\underline{c}, \bar{c}\right)$ and firms choose a mixed-strategy probability p of innovating. This mixed strategy equilibrium is a tractable representation of an environment in which duopolists make smoothly-varying choices over the probability of successful stochastic innovation.

1.3.1 Symmetric Actions: Uncoordinated Equilibria versus Joint Profit Maximization

Suppose that instead of an uncoordinated mixed-strategy $p \in [0,1]$ chosen by firms, the duopolists coordinate on a symmetric $p_J \in [0,1]$ to maximize joint profits. When $S > 0$, the joint maximization arrangement includes knowledge transfer. In this case, I show in Appendix Section 1.9.2 that firms will choose p_J as follows:

$$p_J = \max\left\{1 - \frac{c}{2(\pi(1,1) - \pi(0,0))}, 0\right\}.$$

When $\theta = 1$, $p \geq p_J$ for all σ and all c; there is overinvestment in the uncoordinated equilibrium.

At the other extreme of $\theta = 0$, I show that may be possible for the relationship $p < p_J$ to hold for values of $c > \underline{c}$. This would correspond to underinvestment in the uncoordinated equilibrium.

1.4 How Innovation Varies with Substitutability and Bargaining Power

In what follows, I analyze the mixed-strategy equilibrium probability of innovation p and show that bargaining power θ alters the slope of the relationship between p and substitutability σ.

1.4.1 Baseline when $\theta = 0$

When $\theta = 0$, the probability of innovation is given by

$$p|_{\theta=0} = \frac{\pi(1,0) - \pi(0,0) - c}{\pi(1,0) - \pi(0,0) - (\pi(1,1) - \pi(0,1)) + \bar{S}}.$$

The probability of innovation p is determined by two factors that are expressed in its numerator and denominator, respectively. The first is the gain to pulling ahead, which is the numerator $\pi(1,0) - \pi(0,0) - c$. The greater this gain, the more each firm innovates. The second factor is the strategic substitutes effect, which is expressed in negated form in the denominator as $\pi(1,0) - \pi(0,0) - (\pi(1,1) - \pi(0,1)) + \bar{S}$. Recall that the strategic substitutes effect represents the innovation-disincentivizing effect that a rival's innovation exerts. The more negative the strategic substitutes effect is, the more positive the denominator, and the less effort that firms put into innovating.

I first analyze how the strategic substitutes effect varies with σ. An imitator reaps all of the knowledge transfer surplus $\bar{S}$ when $\theta = 0$, and when $\sigma = 0$ this surplus is large. The result is that a rival's innovation negatively affects a firm's probability of innovating, because the firm would rather purchase the innovation from its rival on very favorable terms. However, holding fixed $\theta = 0$, surplus decreases in σ and therefore the incentive to free-ride diminishes in σ as well. As long as $\bar{S} > 0$, this diminishing free-riding effect makes the strategic substitutes effect less negative, causes the denominator of p to decrease in σ, and increases p.[5]

Assume furthermore that the gain to pulling ahead, $\pi(1,0) - \pi(0,0) - c$, is weakly increasing in σ:

[5] If S is expanded in the denominator and terms are cancelled, the denominator is equivalent to $\pi(1,1) - \pi(0,0)$ which is decreasing in σ.

$$\frac{d\{\pi(1,0)-\pi(0,0)\}}{d\sigma} \geq 0.$$

I show in Section 1.6 that this assumption is valid to a first approximation in an environment of nested CES, price competition, and constant marginal cost-reducing innovation.

As σ increases, the gain to pulling ahead and the strategic substitutes effect both contribute to an increase in p:

Proposition 2. *When innovator bargaining power $\theta = 0$ and $\bar{S} > 0$, the mixed-strategy equilibrium p is increasing in σ:*

$$\frac{dp}{d\sigma} > 0.$$

1.4.2 Varying θ

As θ increases from 0, there is an extra incentive to innovate to sell the technology. However, holding fixed a $\theta > 0$, this extra innovation incentive decreases in σ because surplus decreases in σ. To see the contribution of $\bar{S}$ analytically, examine the cross derivative:

$$\frac{d^2p}{d\sigma d\theta} = \frac{\frac{d\bar{S}}{d\sigma}}{(\pi(1,1)-\pi(0,0))} - \frac{\bar{S}\frac{d(\pi(1,1)-\pi(0,0))}{d\sigma}}{(\pi(1,1)-\pi(0,0))^2}.$$

The first term, $\frac{\frac{d\bar{S}}{d\sigma}}{(\pi(1,1)-\pi(0,0))}$, is negative due to $\frac{d\bar{S}}{d\sigma} < 0$, and represents the decreasing incentive to innovate and sell technology as σ increases. Once the surplus terms in this expression are expanded and common terms cancelled, the negative contribution of $\frac{d\bar{S}}{d\sigma}$ prevails.

In Appendix Section 1.9.3 I prove the following proposition:

Proposition 3. *Let $\bar{\sigma}$ be the highest σ for which $S \geq 0$. On $\sigma \in [0,\bar{\sigma})$, the slope of p with respect to σ is decreasing as bargaining power θ increases:*

$$\frac{d^2p}{d\sigma d\theta}\Big|_{\sigma\in[0,\bar{\sigma})} \leq 0,$$

and there exists a $\underline{\sigma} \in (0, \bar{\sigma})$ *such that*

$$\frac{d^2 p}{d\sigma d\theta}|_{\sigma \in (\underline{\sigma}, \bar{\sigma})} < 0.$$

The relationship of p to substitutability and bargaining power can be seen in Figure 1.3. An increase in θ for a given σ will increase innovation. However, the same increase in θ also decreases the slope of innovation with respect to substitutability.

Despite this result on how bargaining power θ affects the slope of p with respect to σ, I show in Appendix Section 1.9.4 that is not possible to determine the sign of the slope $\frac{dp}{d\sigma}$ at $\theta \gg 0$. In a similar vein, I show in Appendix Section 1.9.5 that the sign of the slope of p in the region $S < 0$ cannot be determined without further assumptions on profit.

1.5 Knowledge Transfer

Any empirical measure of the unconditional incidence of knowledge transfer implicitly incorporates both innovation and the conditional probability of subsequent knowledge transfer. The empirical frequency corresponds to the ex-ante probability of knowledge transfer q in the model. Such an object q reflects the probability that, in the course of the game, an asymmetric innovation outcome will occur and the innovation is subsequently transferred from innovator to imitator. The former is given by $2p(1-p)$, and the latter is given by $I_{\bar{S}>0}$. Together, these two factors constitute the expression for q:

$$q = 2p(1-p) \cdot I_{\bar{S}>0}.$$

The incidence of knowledge transfer follows a rising-then-falling pattern in σ, when $\theta = 0$ and the cost of innovation c is sufficiently high that $p|_{\sigma=0} \ll \frac{1}{2}$. Innovation p is initially increasing in σ as would-be imitators substitute away from knowledge transfer toward their own innovation. New innovation generates opportunities for knowledge transfer $2p(1-p)$, which also rise in σ. However, at a sufficiently high degree of substitutability, surplus becomes negative and conditional knowledge transfer ceases: $I_{\bar{S}>0} = 0$. This creates the falling portion of the pattern. This rising-then-falling relationship is depicted in Figure 1.4 for $\theta = 0$.

The model only makes a sharp prediction for this rising-then-falling pattern for values of θ close to and including 0, since it is only at $\theta = 0$ that the sign of the slope $\frac{dp}{d\sigma}$ can be established. It is also possible for a monotonically decreasing relationship of q to σ to exist. For example, suppose that $\theta = 0$, and the innovation cost c is sufficiently low to result in $p|_{\sigma=0} \gg \frac{1}{2}$. Then q may be decreasing in σ over the entire subset of the domain in which surplus is positive. Where substitutability is high and surplus is negative, $q = 0$. Putting these two parts of the domain together, knowledge transfer would be monotonically decreasing in σ.

1.6 Correspondence of Profit Function to Actual Demand System

In the preceding sections, the analysis was predicated on a profit function $\pi(H_i, H_{-i})$ that I characterized with reduced-form assumptions. In this section I show that those reduced-form assumptions are consistent with an environment of nested CES demand, price competition, and constant marginal cost-reducing innovation. In particular, the nested CES demand is summarized by two parameters: an outer industry elasticity ν and an inner elasticity of substitution ρ between duopolists' products, where $\rho \geq \nu$. A substitutability index $\sigma \in [0,1]$ can be constructed by varying $\frac{\rho-1}{\rho}$ and normalizing as follows:

$$\sigma = \frac{\frac{\rho-1}{\rho} - \frac{\nu-1}{\nu}}{1 - \frac{\nu-1}{\nu}}.$$

A technology level of $H_i = 0$ corresponds to a normalized constant marginal cost of 1. Innovation of $H_i = 1$ is modeled as reducing constant marginal cost of firm i to $\frac{1}{\gamma} < 1$.

1.6.1 Profit Rankings over Technology Outcomes

Profits are a function of the possible technology outcomes that can be realized in this game:

$$\pi\left(H_i, H_{-i}\right) \in \left\{\pi\left(0,0\right), \pi\left(0,1\right), \pi\left(1,0\right), \pi\left(1,1\right)\right\}.$$

In Appendix Section 1.9.6, I prove that profits in the specified competitive environment can be multiplicatively separated into relative and absolute technological components:

$$\pi\left(H_i, H_{-i}; \sigma, \nu, \gamma\right) = \kappa\left(H_i - H_{-i}; \sigma, \nu, \gamma\right) \cdot \gamma^{H_i \cdot (\nu - 1)}.$$

I also show that $\kappa\left(H_i - H_{-i}; \sigma\right)$ is continuous in σ.

Note that $H_i - H_{-i} \in \{-1, 0, 1\}$. In the Appendix Section 1.9.7 I prove that $\kappa\left(1\right) \geq \kappa\left(0\right) \geq \kappa\left(-1\right)$. That leads to the following proposition:

Proposition 4. *In an environment of nested CES demand with outer and inner elasticities of substitution ν and ρ, normalized index of substitutability $\sigma = \frac{\frac{\rho-1}{\rho} - \frac{\nu-1}{\nu}}{1 - \frac{\nu-1}{\nu}}$, price competition, and innovation that reduces constant marginal cost by a factor $\gamma > 1$, the following profit relationships hold:*[6]

$$\pi\left(1,0\right) \geq \pi\left(1,1\right) \geq \pi\left(0,0\right) \geq \pi\left(0,1\right), \forall \sigma \in [0,1].$$

1.6.2 Rents to Level Innovation

The rents to level innovation, $\pi\left(1,1\right) - \pi\left(0,0\right)$ can be written as follows:

$$\pi\left(1,1\right) - \pi\left(0,0\right) = \kappa\left(0\right)\left(\gamma^{\nu-1} - 1\right)$$

To understand the rents to level innovation, it suffices to characterize $\kappa(0)$. In Appendix Section 1.9.8, I prove the following:

[6] The relationship between $\pi\left(1,1\right)$ and $\pi\left(0,0\right)$ is a weak rather than strong inequality because, as I show in the next subsection, $\kappa\left(0\right)|_{\sigma=1} = 0$.

$$\begin{aligned}\kappa(0)\,|_{\sigma=0} &> 0, \\ &\text{and} \\ \kappa(0)\,|_{\sigma=1} &= 0.\end{aligned}$$

I also prove that the slope is negative:

$$\frac{d\kappa(0)}{d\sigma} < 0.$$

These results for $\kappa(0)$ constitute a proof of the following proposition.

Proposition 5. *The rents to level innovation* $\pi(1,1) - \pi(0,0)$ *have the following properties:*

$$\begin{aligned}\{\pi(1,1) - \pi(0,0)\}\,|_{\sigma=0} &> 0, \\ \{\pi(1,1) - \pi(0,0)\}\,|_{\sigma=1} &= 0, \\ \frac{d\{\pi(1,1) - \pi(0,0)\}}{d\sigma} &< 0.\end{aligned}$$

1.6.3 Gains to Pulling Ahead and Gains to Not Falling Behind

The gain to not falling behind is defined as $\pi(1,1) - \pi(0,1)$ and the gain to pulling ahead is defined as $\pi(1,0) - \pi(0,0)$. Previously, I defined the difference of these two gains $\{\pi(1,1) - \pi(0,1)\} - \{\pi(1,0) - \pi(0,0)\}$ as the strategic substitutes effect that a rival's innovation has on a firm's own innovation incentives.

In Appendix Section 1.9.9 I prove that at the endpoints of $\sigma = 0, 1$ the strategic substitutes effect is equal to zero and negative, respectively:

$$\begin{aligned}\{(\pi(1,1)-\pi(0,1))-(\pi(1,0)-\pi(0,0))\}|_{\sigma=0} &= 0,\\ \{(\pi(1,1)-\pi(0,1))-(\pi(1,0)-\pi(0,0))\}|_{\sigma=1} &< 0.\end{aligned}$$

To characterize the strategic substitutes effect in the interval $\sigma \in (0,1)$, I numerically compute profit at 1600 points in the range $\frac{\nu-1}{\nu} \in [0.1, 0.5], \gamma \in [1.01, 1.40]$. A representative set of results with outer nested CES parameter value $\frac{\nu-1}{\nu} = 0.29$ and growth factor $\gamma = 1.14$ are plotted in Figure 1.5. Here it is evident that the following relationship holds:

$$\{(\pi(1,1)-\pi(0,1))-(\pi(1,0)-\pi(0,0))\} \leq 0, \forall\sigma \in [0,1].$$

Furthermore, a stronger result holds in all of the numerical results:

$$\frac{d\{(\pi(1,1)-\pi(0,1))-(\pi(1,0)-\pi(0,0))\}}{d\sigma} \leq 0, \forall\sigma \in [0,1].$$

Figure 1.5 does highlight a complication to the stylized assumptions, which is that the gain to pulling ahead, $\pi(1,0)-\pi(0,0)$, is not monotonically increasing in σ. Instead, for low values of σ, there is an initially decreasing segment of $\pi(1,0)-\pi(0,0)$. In the aforementioned numerical results, this non-monotonic feature is always present.[7] In particular, the inflection point at which the slope $\frac{d(\pi(1,0)-\pi(0,0))}{d\sigma}$ turns positive varies over $\sigma \in [0.03, 0.18]$. The depth of the decreasing segment, as a fraction of the total range of $\pi(1,0)-\pi(0,0)$, varies in $[0.002, 0.01]$. In these numerical results, $\pi(1,0)-\pi(0,0)$ is always weakly concave. Thus, the stylized assumption of $\frac{d(\pi(1,0)-\pi(0,0))}{d\sigma} \geq 0$ accurately describes the behavior of $\pi(1,0)-\pi(0,0)$ over most of $\sigma \in [0,1]$.

How does the non-monotonicity impact the two-stage model results for innovation p and knowledge transfer incidence q? I proceed with the parameter values $\frac{\nu-1}{\nu} = 0.29$ and $\gamma = 1.14$ from above, and plot p and q in Figure 1.6 for $\theta = 0$ as would result in the two-stage model.

[7] It turns out that an increase in the outer CES parameter $\frac{\nu-1}{\nu}$ has a first order impact on increasing the length and depth of the initially decreasing segment, while γ has a lesser effect.

Qualitatively, the incidence of knowledge transfer q can still be described in a first-order manner as first rising, then falling. It is also the case that if the cost of innovation c is adjusted high enough, then the complications to p and q disappear as both curves are shifted downward and truncated at 0, and they behave exactly as characterized in the two-stage model.

1.7 Conclusion

In the context of technology licensing or inventor job transitions, knowledge diffusion between firms is the outcome of decisions to innovate and then allow or promote the transfer of that innovation. These voluntary transfers of knowledge depend upon both the degree of competition between firms and the extent of appropriability of innovations.

In a two-stage model, I demonstrate that bargaining power determines how both innovation and knowledge transfer vary with product substitutability. In particular, I show that a non-monotonic relationship of knowledge transfer to product substitutability can arise when the imitator reaps most of the surplus from knowledge transfer. As product substitutability increases, the gain to pulling ahead rises while the free-riding benefit from knowledge transfer falls, prompting a would-be imitator to innovate more. This increases the probability of asymmetric outcomes, which entail knowledge transfer. However, surplus eventually becomes negative as product substitutability continues to increase. The result is that knowledge transfer ceases, thus producing the falling part of the rising-then-falling non-monotonicity.

1.8 Tables and Figures

Figure 1.1: Timing of Two-Stage Game

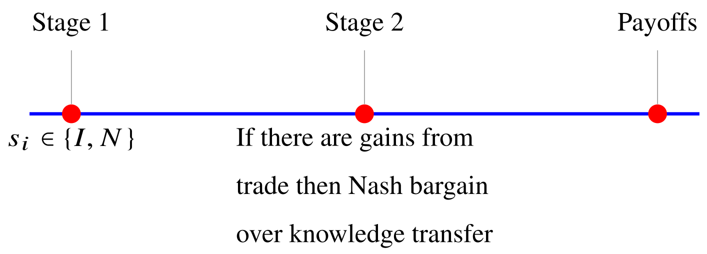

Figure 1.2: Symmetric Equilibria as a Function of Innovation Cost

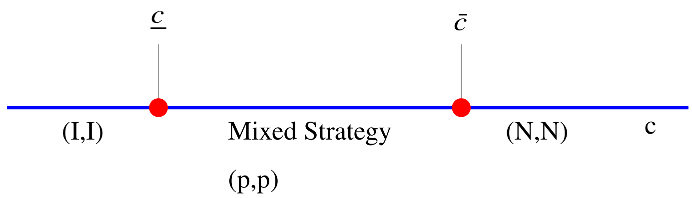

Figure 1.3: Innovation Probability

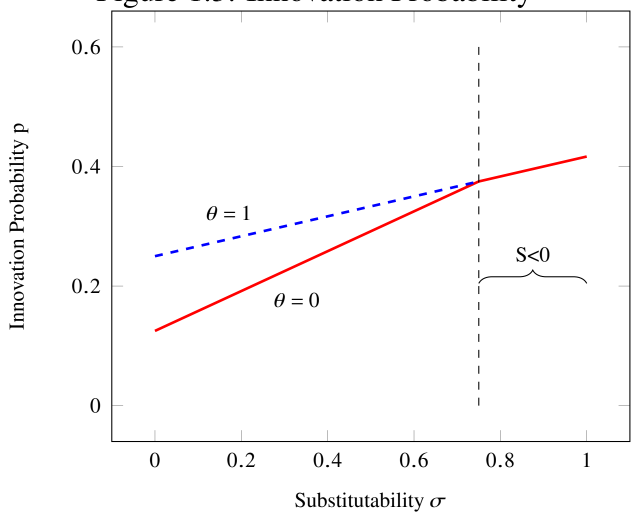

Figure 1.4: Pattern of Knowledge Transfer for $\theta = 0$

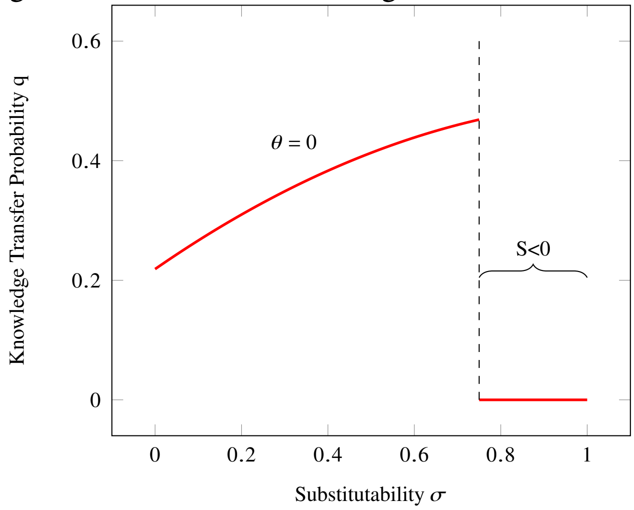

Figure 1.5: Gain to Pulling Ahead vs. Not Falling Behind

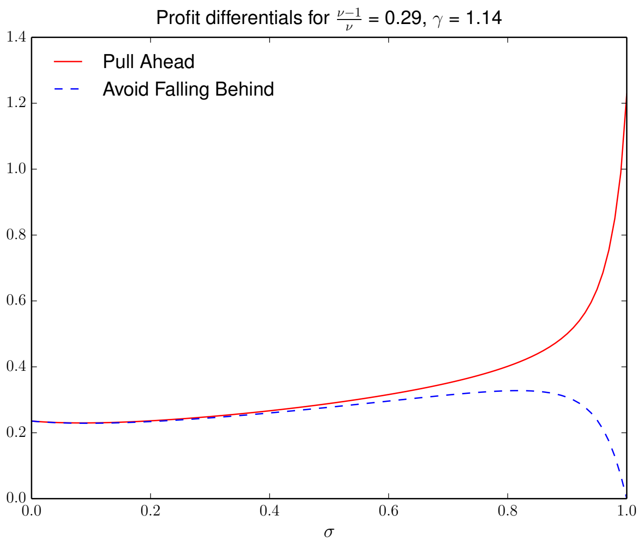

Figure 1.6: Game Actions with Nested CES Demand

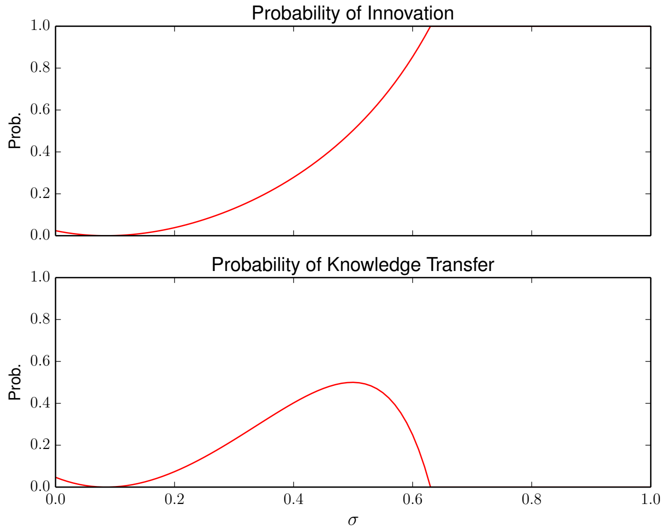

Table 1.1: Game Payoffs, First Stage Perspective

	I	N
I	$(\pi(1,1)-c, \pi(1,1)-c)$	$\left(\pi(1,0)+\theta\bar{S}-c, \pi(0,1)+(1-\theta)\bar{S}\right)$
N	$\left(\pi(0,1)+(1-\theta)\bar{S}, \pi(1,0)+\theta\bar{S}-c\right)$	$(\pi(0,0), \pi(0,0))$

1.9 Appendices

1.9.1 Appendix 1: Surplus

Recall that surplus S can be written as:

$$S = \{\pi(1,1) - \pi(0,1)\} - \{\pi(1,0) - \pi(0,0)\} + \{\pi(1,1) - \pi(0,0)\}.$$

Recall that both the strategic substitutes effect and the rents to level innovation are decreasing in σ:

$$\begin{aligned}
\frac{d\{\pi(1,1) - \pi(0,1) - (\pi(1,0) - \pi(0,0))\}}{d\sigma} &\leq 0, \\
\frac{d\{\pi(1,1) - \pi(0,0)\}}{d\sigma} &< 0.
\end{aligned}$$

Putting these together,

$$\frac{dS}{d\sigma} = \frac{d\{\pi(1,1) - \pi(0,1) - (\pi(1,0) - \pi(0,0))\}}{d\sigma} + \frac{d\{\pi(1,1) - \pi(0,0)\}}{d\sigma} < 0.$$

To prove that surplus is positive at $\sigma = 0$:

Recall that at $\sigma = 0$,

$$\begin{aligned}
\{\pi(1,1) - \pi(0,1) - (\pi(1,0) - \pi(0,0))\}|_{\sigma=0} &= 0, \\
\{\pi(1,1) - \pi(0,0)\}|_{\sigma=0} &> 0.
\end{aligned}$$

Putting these together,

$$S|_{\sigma=0} = \{\pi(1,1) - \pi(0,0)\}|_{\sigma=0} > 0.$$

Finally, to prove that surplus is negative at $\sigma = 1$:

It was assumed that

$$\begin{aligned} \{\pi(1,1) - \pi(0,1) - (\pi(1,0) - \pi(0,0))\}|_{\sigma=1} &< 0, \\ \{\pi(1,1) - \pi(0,0)\}|_{\sigma=1} &= 0. \end{aligned}$$

Putting these together,

$$S|_{\sigma=1} = \{\pi(1,1) - \pi(0,1) - (\pi(1,0) - \pi(0,0))\}|_{\sigma=1} < 0.$$

In summary, I have proven that surplus demonstrates the following properties:

$$\begin{aligned} S|_{\sigma=0} &> 0, \\ S|_{\sigma=1} &< 0, \\ \frac{dS}{d\sigma} &< 0. \end{aligned}$$

1.9.2 Appendix 2: Uncoordinated vs. Coordinated Innovation

Here I compare the randomizing case to a jointly maximized, symmetric case.

Suppose $S > 0$. The joint maximization problem is as follows:

$$\begin{aligned} \max_{p_J} \quad & p_J^2 \cdot 2(\pi(1,1) - c) + (1 - p_J)^2 2\pi(0,0) + 2p_J(1 - p_J)(2\pi(1,1) - c) \\ s.t. \quad & \\ p_J \in & \; [0,1]. \end{aligned}$$

Let p_J^* denote the uncensored p_J determined by the first order conditions when ignoring the probability constraints:

$$\begin{aligned}
2 \cdot p_J^* \cdot 2(\pi(1,1) - c) & \\
-2\left(1 - p_J^*\right) 2\pi(0,0) & \\
+2\left(1 - 2p_J^*\right)(2\pi(1,1) - c) &= 0 \\
p_J^* &= \frac{2(\pi(1,1) - \pi(0,0)) - c}{2(\pi(1,1) - \pi(0,0))}.
\end{aligned}$$

Compare to the uncoordinated symmetric p when $S > 0$:

$$\begin{aligned}
p &= \frac{\pi(1,0) - \pi(0,0) + \theta\bar{S} - c}{\pi(1,0) - \pi(0,0) - (\pi(1,1) - \pi(0,1)) + \bar{S}} \\
&= \frac{\pi(1,0) - \pi(0,0) + \theta\bar{S} - c}{\pi(1,1) - \pi(0,0)}.
\end{aligned}$$

When $\theta = 1$:

The difference $p - p_J^*$ is a linear function of innovation cost c:

$$p - p_J^* = \frac{2(\pi(1,1) - \pi(0,1)) - c}{2(\pi(1,1) - \pi(0,0))}.$$

At $c = \bar{c}$ where $p = 0$,

$$p - p_J^* = \frac{\pi(0,0) - \pi(0,1)}{2(\pi(1,1) - \pi(0,0))} \geq 0.$$

At $c = \underline{c}$ where $p = 1$,

$$p - p_J^* = \frac{\pi(1,1) - \pi(0,0) + (\pi(0,0) - \pi(0,1))}{2(\pi(1,1) - \pi(0,0))} \geq \frac{1}{2}.$$

These results demonstrate that for $\theta = 1$, overinvestment in the uncoordinated equilibrium exists on the entire interval $c \in [\underline{c}, \bar{c})$, as well as for a range of values $c < \underline{c}$.

Now, suppose $\theta = 0$.

The difference $p - p_J^*$ is still a linear function of innovation cost c:

$$p - p_J^* = \frac{2(\pi(1,0) - \pi(1,1)) - c}{2(\pi(1,1) - \pi(0,0))}$$

At $c = \bar{c}$ where $p = 0$,

$$\begin{aligned} p - p_J^* &= \frac{\pi(1,0) - \pi(1,1) - (\pi(1,1) - \pi(0,0))}{2(\pi(1,1) - \pi(0,0))} \\ &= \frac{-S + \pi(0,0) - \pi(0,1)}{2(\pi(1,1) - \pi(0,0))} \end{aligned}$$

It is not possible to determine the sign of this expression. In particular, it may be the case that $(p - p_J)\,|_{c=\bar{c}} < 0$.

At $c = \underline{c}$ where $p = 1$,

$$p - p_J^* = \frac{\pi(1,0) - \pi(1,1)}{2(\pi(1,1) - \pi(0,0))} \geq 0.$$

Overall, the possibility for underinvestment in the uncoordinated equilibrium exists for $c > \underline{c}$.

1.9.3 Appendix 3: Proof that the Slope of Innovation versus Competition is Decreasing in θ for Positive Surplus

Let $\bar{\sigma}$ be the highest σ at which surplus S is still weakly positive. We know that $S|_{\sigma=\bar{\sigma}} = 0$ and $\bar{\sigma} < 1$ because $S|_{\sigma=1} < 0$ and $\pi(H_i, H_{-i}; \sigma)$ is continuous in σ. Then over the interval $\sigma \in [0, \bar{\sigma})$, the cross-derivative of p with respect to σ and θ can be written as follows:

$$
\begin{aligned}
\frac{d^2p}{d\sigma d\theta}|_{\sigma\in[0,\bar{\sigma})} &= \frac{(\pi(1,1)-\pi(0,0))\frac{d\bar{S}}{d\sigma}-\bar{S}\frac{d(\pi(1,1)-\pi(0,0))}{d\sigma}}{(\pi(1,1)-\pi(0,0))^2} \\
&= \frac{(\pi(1,1)-\pi(0,0))\frac{d(\pi(1,1)-\pi(0,1)-(\pi(1,0)-\pi(0,0)))}{d\sigma}}{(\pi(1,1)-\pi(0,0))^2} \\
&\quad +\frac{((\pi(1,0)-\pi(0,0))-(\pi(1,1)-\pi(0,1)))\frac{d(\pi(1,1)-\pi(0,0))}{d\sigma}}{(\pi(1,1)-\pi(0,0))^2}.
\end{aligned}
$$

Examining the numerator of the first term, it is true that

$$(\pi(1,1)-\pi(0,0))\frac{d(\pi(1,1)-\pi(0,1)-(\pi(1,0)-\pi(0,0)))}{d\sigma}\leq 0.$$

Examining the numerator of the second term, it is true that

$$((\pi(1,0)-\pi(0,0))-(\pi(1,1)-\pi(0,1)))\frac{d(\pi(1,1)-\pi(0,0))}{d\sigma}\leq 0.$$

Putting these together,

$$\frac{d^2p}{d\sigma d\theta}|_{\sigma\in[0,\bar{\sigma})}\leq 0.$$

However, a stronger result can be proven.

Because of the assumptions made on $\pi(1,1)-\pi(0,0)$, it is true that $(\pi(1,1)-\pi(0,0))|_{\sigma=\bar{\sigma}}>0$. Thus, it is also true that $((\pi(1,1)-\pi(0,1))-(\pi(1,0)-\pi(0,0)))|_{\sigma=\bar{\sigma}}<0$. By continuity of $\pi(H_i,H_{-i};\sigma)$, $\exists\underline{\sigma}<\bar{\sigma}$ such that the strategic substitutes effect has the following properties:

$$
\begin{aligned}
((\pi(1,1)-\pi(0,1))-(\pi(1,0)-\pi(0,0)))|_{\sigma>\underline{\sigma}} &< 0, \\
((\pi(1,1)-\pi(0,1))-(\pi(1,0)-\pi(0,0)))|_{\sigma\leq\underline{\sigma}} &\geq 0.
\end{aligned}
$$

This result is important because the second term of $\frac{\partial^2 p}{\partial\sigma\partial\theta}$ is then strictly negative on $\sigma\in$

$\left(\underline{\sigma}, \bar{\sigma}\right)$, which yields the stronger result below:

$$\frac{d^2 p}{d\sigma d\theta}|_{\sigma \in \left(\underline{\sigma}, \bar{\sigma}\right)} < 0.$$

1.9.4 Appendix 4: The Slope of p When $\theta = 1$

When $\theta = 1$, the slope of p in σ is given by:

$$\begin{aligned} \frac{dp}{d\sigma}|_{\theta=1} &= \frac{\left(\pi\left(1,1\right) - \pi\left(0,0\right)\right) \frac{d(\pi(1,1)-\pi(0,1))}{d\sigma}}{\left(\pi\left(1,1\right) - \pi\left(0,0\right)\right)^2} \\ &\quad + \frac{\left(c - \left(\pi\left(1,1\right) - \pi\left(0,1\right)\right)\right) \frac{d(\pi(1,1)-\pi(0,0))}{d\sigma}}{\left(\pi\left(1,1\right) - \pi\left(0,0\right)\right)^2} \end{aligned}$$

The second term is negative. However, the sign of $\frac{d(\pi(1,1)-\pi(0,1))}{d\sigma}$ in the first term may be positive or negative, so it is not possible to know the sign of the slope $\frac{dp}{d\sigma}$.

1.9.5 Appendix 5: The Slope of p When $\bar{S} = 0$

When $\bar{S} = 0$, innovation p is given by:

$$p = \frac{\pi\left(1,0\right) - \pi\left(0,0\right) - c}{\pi\left(1,0\right) - \pi\left(0,0\right) - \left(\pi\left(1,1\right) - \pi\left(0,1\right)\right)}$$

The slope can be written as being proportional to the following two terms:

$$\begin{aligned} \frac{dp}{d\sigma} &\propto \left(c - \left(\pi\left(1,1\right) - \pi\left(0,1\right)\right)\right) \frac{d\left(\pi\left(1,0\right) - \pi\left(0,0\right)\right)}{d\sigma} \\ &\quad + \left(\pi\left(1,0\right) - \pi\left(0,0\right) - c\right) \frac{d\left(\pi\left(1,1\right) - \pi\left(0,1\right)\right)}{d\sigma}. \end{aligned}$$

The first term is weakly positive. Without assumptions on the sign of $\frac{d(\pi(1,1)-\pi(0,1))}{d\sigma}$ in the second term, it is not possible to know what the sign of $\frac{dp}{d\sigma}$ is.

However, I show here that it is possible to bound the slope from above when $S < 0$ with the innovation that would occur in a hypothetical situation in which knowledge transfers still continue and $\theta = 0$. Let $\frac{\hat{dp}}{d\sigma}$ denote the slope of this hypothetical innovation.

Taking their difference,

$$\begin{aligned}\frac{\hat{dp}}{d\sigma} - \frac{dp}{d\sigma}|_{\bar{S}=0} \quad = \quad & \frac{-\frac{d(\pi(1,0)-\pi(0,0))}{d\sigma} S}{(\pi(1,1) - \pi(0,0))(\pi(1,0) - \pi(0,0) - (\pi(1,1) - \pi(0,1)))} \\ & - (\pi(1,0) - \pi(0,0) - c) \\ & \times \left(\frac{\frac{d(\pi(1,1)-\pi(0,0))}{d\sigma}}{(\pi(1,1) - \pi(0,0))^2} - \frac{\frac{d((\pi(1,0)-\pi(0,0))-(\pi(1,1)-\pi(0,1)))}{d\sigma}}{((\pi(1,0) - \pi(0,0)) - (\pi(1,1) - \pi(0,1)))^2} \right)\end{aligned}$$

The first term is weakly positive and the second term is strictly positive. Thus the following relationship holds:

$$\frac{\hat{dp}}{d\sigma} > \frac{dp}{d\sigma}|_{\bar{S}=0}.$$

1.9.6 Appendix 6: Derivation of Profit for Nested CES Demand

Let demand for the duopoly's aggregated output Q at aggregated price index P be subject to CES demand with outer elasticity of substitution ν and scale demand factor A:

$$Q = AP^{-\nu}$$

The duopoly's aggregated output Q is given as a CES function of individual outputs q_i, q_{-i} with inner elasticity of substitution $\rho \geq \nu$:

$$Q = \left(q_i^{\frac{\rho-1}{\rho}} + q_{-i}^{\frac{\rho-1}{\rho}} \right)^{\frac{\rho}{\rho-1}}$$

The duopoly's aggregate price index is given by

$$P = \left(p_i^{1-\rho} + p_{-i}^{1-\rho} \right)^{\frac{1}{1-\rho}}$$

Demand q_i is given by

$$\begin{aligned} q_i &= Q\left(\frac{p_i}{P}\right)^{-\rho} \\ &= AP^{\rho-\nu}p_i^{-\rho} \end{aligned}$$

In each period, firms solve a static price competition problem taking their opponent's price p_{-i} as given, with constant marginal cost ψ_i:

$$\begin{aligned} \max_{p_i} & \quad (p_i - \psi_i)\, q_i \\ s.t. & \\ q_i &= AP^{\rho-\nu}p_i^{-\rho} \end{aligned}$$

Defining an elasticity e_i, the optimal price p_i is given by the familiar expression:

$$p_i = \frac{e_i}{e_i - 1}\psi_i$$

The elasticity in turn is given by an expression that depends upon p_i:

$$e_i = \rho - (\rho - \nu)\left(\frac{p_i}{P}\right)^{1-\rho}$$

Profit π_i can be written as:

$$
\begin{aligned}
\pi_i &= \frac{1}{e_i - 1} q_i \psi_i \\
&= \frac{1}{e_i - 1} \frac{1}{p_i} A P^{1-\nu} \left(\frac{p_i}{P}\right)^{1-\rho} \psi_i \\
&= \frac{1}{\rho - (\rho - \nu)\left(\frac{p_i}{P}\right)^{1-\rho}} A P^{1-\nu} \left(\frac{p_i}{P}\right)^{1-\rho}
\end{aligned}
$$

Define a new variable $R_i = \left(\frac{p_{-i}}{p_i}\right)^{\rho-1}$.

Using R_i, I can re-write price-related objects in terms of R_i:

$$
\begin{aligned}
\left(\frac{p_i}{P}\right)^{1-\rho} &= \frac{1}{1 + \frac{1}{R_i}} \\
&= \frac{R_i}{R_i + 1}
\end{aligned}
$$

and

$$
\begin{aligned}
p_i &= \frac{\rho - (\rho - \nu)\left(\frac{p_i}{P}\right)^{1-\rho}}{\rho - (\rho - \nu)\left(\frac{p_i}{P}\right)^{1-\rho} - 1} \psi_i \\
&= \frac{\rho + \nu R_i}{\rho - 1 + (\nu - 1) R_i} \psi_i
\end{aligned}
$$

and finally

$$
\begin{aligned}
P^{1-\nu} &= \left(p_i \left(1 + \frac{1}{R_i}\right)^{\frac{1}{1-\rho}}\right)^{1-\nu} \\
&= \left(\frac{\rho + \nu R_i}{\rho - 1 + (\nu - 1) R_i} \psi_i \left(1 + \frac{1}{R_i}\right)^{\frac{1}{1-\rho}}\right)^{1-\nu}.
\end{aligned}
$$

Replace price-related terms in π_i with functions of R_i:

$$
\begin{aligned}
\pi_i &= \frac{1}{\rho-(\rho-\nu)\frac{R_i}{R_i+1}} A\left(\frac{\rho+\nu R_i}{\rho-1+(\nu-1)R_i}\psi_i\left(1+\frac{1}{R_i}\right)^{\frac{1}{1-\rho}}\right)^{1-\nu}\frac{R_i}{R_i+1} \\
&= \frac{R_i}{\rho+\nu R_i} A\left(\frac{\rho-1+(\nu-1)R_i}{\rho+\nu R_i}\left(1+\frac{1}{R_i}\right)^{\frac{1}{\rho-1}}\right)^{\nu-1}\left(\frac{1}{\psi_i}\right)^{\nu-1} \\
&= \phi(R_i)\left(\frac{1}{\psi_i}\right)^{\nu-1}
\end{aligned}
$$

I define a fixed point in R_i as follows:

$$
\begin{aligned}
R_i &= \left(\frac{p_{-i}}{p_i}\right)^{\rho-1} \\
R_i &= \left(\frac{\frac{\rho R_i+\nu}{(\rho-1)R_i+\nu-1}\psi_{-i}}{\frac{\rho\frac{1}{R_i}+\nu}{(\rho-1)\frac{1}{R_i}+\nu-1}\psi_i}\right)^{\rho-1} \\
R_i &= \left(\frac{\rho R_i+\nu}{(\rho-1)R_i+(\nu-1)}\frac{(\nu-1)R_i+(\rho-1)}{\nu R_i+\rho}\right)^{\rho-1}\left(\frac{\psi_{-i}}{\psi_i}\right)^{\rho-1}.
\end{aligned}
$$

In the two-stage model, the marginal cost ψ_i is equal to 1 if $H_i = 0$, and is equal to $\frac{1}{\gamma} < 1$ if $H_i = 1$. R_i can be re-written as a function of $H_i - H_{-i}$ as follows:

$$
R(H_i - H_{-i}) = \left(\frac{\rho R_i+\nu}{(\rho-1)R_i+(\nu-1)}\frac{(\nu-1)R_i+(\rho-1)}{\nu R_i+\rho}\right)^{\rho-1}\left(\gamma^{H_i-H_{-i}}\right)^{\rho-1}.
$$

See Appendix Section 1.9.7 for a demonstration that the fixed point solution $R(H_i - H_{-i};\rho)$ is continuous as a function of ρ. I also show that as $\rho\uparrow\infty$, the ratio $\frac{R(1)}{\rho}$ converges to a positive constant $\frac{\gamma-1}{\gamma-\nu(\gamma-1)}$. This ensures that the function $\phi(R(H_i, H_{-i};\rho);\rho)$ is a continuous function of σ on $\sigma\in[0,1]$. In particular, at $\sigma = 1$, which corresponds to $\rho\uparrow\infty$, $\lim_{\rho\uparrow\infty}\phi(R(1)) = A\gamma^{-\nu}(\gamma-1)$. This corresponds to the Bertrand competition solution.

We can then re-write $\phi(R_i) = \phi(R(H_i - H_{-i})) = \kappa(H_i - H_{-i})$. $\kappa(H_i - H_{-i})$ is a continuous function of σ.

Finally, profit π_i can be written as follows:

$$\pi_i = \kappa(H_i - H_{-i})\gamma^{H_i \cdot (\nu-1)}.$$

π_i is a continuous function of σ.

1.9.7 Appendix 7: Proof that Profit Increases in Relative Technological Position

Here I prove that $\frac{d\phi(R)}{dR} \geq 0$ by proving $\frac{dlog(\phi(R))}{dR} \geq 0$.

The slope of the log is given by:

$$\begin{aligned}
\frac{dlog(\phi(R))}{dR} &= \frac{\rho}{R(\rho+\nu R)}(1-\Gamma), \\
\Gamma &= \frac{\left(\frac{\nu-1}{\rho-1}\right)\{(\rho-1)+((\nu-1)+(\rho-1))R\} + \left(1-\frac{\nu}{\rho}\frac{\rho-\nu}{\rho-1}\right)\left\{(\nu-1)R^2\right\}}{\{(\rho-1)+((\nu-1)+(\rho-1))R\}+\{(\nu-1)R^2\}}.
\end{aligned}$$

Note that for $\rho > \nu > 1$, $\Gamma \in (0,1)$. At the endpoint $\rho = \nu > 1$, $\Gamma = 1$.

Thus for $\rho \geq \nu > 1$,

$$\begin{aligned}
\frac{dlog(\phi(R))}{dR} &\in [0, \frac{\rho}{R(\rho+\nu R)}), \\
\frac{d\phi(R)}{dR} &\geq 0.
\end{aligned}$$

Now I show that $R(1) > 1$.

$$R(1) = \left(\frac{\rho R(1) + \nu}{(\rho - 1) R(1) + (\nu - 1)} \frac{(\nu - 1) R(1) + (\rho - 1)}{\nu R(1) + \rho} \right)^{\rho - 1} \gamma^{\rho - 1}$$

Rearrange:

$$\frac{R(1)}{\gamma^{\rho-1}} = \left(\frac{\rho R(1) + \nu}{(\rho - 1) R(1) + (\nu - 1)} \frac{(\nu - 1) R(1) + (\rho - 1)}{\nu R(1) + \rho} \right)^{\rho - 1}$$

First, I show that the right hand side, hereafter abbreviated RHS, has a negative slope everywhere in $R(1)$:

$$\begin{aligned} \frac{dlog(RHS)}{dR} &= \frac{(\rho - 1)(\nu - \rho)}{(\rho R(1) + \nu)((\rho - 1) R(1) + (\nu - 1))} \\ &\quad + \frac{(\rho - 1)(\nu - \rho)}{(\nu R(1) + \rho)((\nu - 1) R(1) + (\rho - 1))} \\ &< 0 \end{aligned}$$

Furthermore, $RHS|_{R(1)=1} = 1$. For any $\gamma > 1$, the left hand side expression $\frac{R(1)}{\gamma^{\rho-1}}$ will intersect the RHS once at $R(1) > 1$.

So the unique solution to the fixed point expression given above is a value $R(1) > 1$. It is also straightforward to see that the solution to the fixed point expression when firms are technologically equal is $R(0) = 1$. Finally, $R(-1) = \frac{1}{R(1)}$. So I have proved the following relationships:

$$R(1) > R(0) = 1 > R(-1) = \frac{1}{R(1)}.$$

Putting together the results on ϕ and R_i, we have the following final result for $\kappa(H_i - H_{-i}) = \phi(R(H_i - H_{-i}))$:

$$\kappa(1) \geq \kappa(0) \geq \kappa(-1).$$

As an aside, for any given $H_i - H_{-i}$, both the LHS and RHS of the R fixed point expression vary continuously in ρ (and in σ), so the solution $R(H_i - H_{-i})$ will vary continuously in ρ and σ as well. In particular, I prove that the limiting behavior of $R(1)$ is well defined as a ratio to ρ, as $\rho \uparrow \infty$ and for $\gamma < \frac{\nu}{\nu-1}$. Start by defining $z = \frac{R(1)}{\rho}$. The fixed point expression can be rewritten as follows:

$$R(1) = \left(\frac{R + \frac{\nu}{\rho}}{\frac{(\rho-1)}{\rho} R + \frac{\nu-1}{\rho}} \frac{(\nu-1)z + \frac{\rho-1}{\rho}}{\nu z + 1} \gamma \right)^{\rho-1}.$$

I first venture a guess that $\lim_{\rho\uparrow\infty} z = \infty$. But this results in a contradiction. The left hand side goes to ∞, whereas the right hand side converges to 0 because $\gamma < \frac{\nu}{\nu-1}$. So the limiting behavior of z must be finite. In that case, the following formulation of the fixed point expression is useful:

$$z^{\frac{1}{\rho-1}} = \left(\frac{R + \frac{\nu}{\rho}}{\frac{(\rho-1)}{\rho} R + \frac{\nu-1}{\rho}} \frac{(\nu-1)z + \frac{\rho-1}{\rho}}{\nu z + 1} \right) \gamma \left(\frac{1}{\rho} \right)^{\frac{1}{\rho-1}}.$$

In this case, the left hand side converges to 1, and the right hand side converges to $\left(\frac{(\nu-1)z+1}{\nu z+1}\right)\gamma$, yielding the following equation:

$$1 = \left(\frac{(\nu-1)z + 1}{\nu z + 1} \right) \gamma.$$

Solving, I find that $\frac{R(1)}{\rho} = \frac{\gamma-1}{\gamma-\nu(\gamma-1)}$ in the limit. Note that for this expression to be finite, it must be the case that the condition $\gamma < \frac{\nu}{\nu-1}$ is fulfilled, which is exactly the assumption I started out with.

1.9.8 Appendix 8: Profit when Both Firms are Technologically Equal

Here I characterize $\kappa(H_i - H_{-i})$ at $H_i - H_{-i} = 0$.

First, evaluate at the endpoints of $\sigma = 0, 1$:

$$
\begin{aligned}
\kappa(0)\,|_{\sigma=0} &= \frac{A}{\nu}\left(\frac{\nu-1}{\nu}\right)^{\nu-1} \\
&> 0,
\end{aligned}
$$

and

$$
\begin{aligned}
\kappa(0)\,|_{\sigma=1} &= A\lim_{\rho=\infty}\frac{1}{\rho+\nu}\lim_{\rho=\infty}\left(\frac{(\rho-1)+(\nu-1)}{\rho+\nu}\right)^{\nu-1}\lim_{\rho=\infty}2^{\frac{\nu-1}{\rho-1}} \\
&= 0.
\end{aligned}
$$

Next, evaluate the slope $\frac{d\kappa(0)}{d\sigma}$:

$$
\begin{aligned}
\frac{d\kappa(0)}{d\sigma} &= \frac{d\phi(R(0;\rho(\sigma));\rho(\sigma))}{d\sigma} \\
&= \frac{\partial\phi(R;\rho)}{\partial\rho}\frac{d\rho}{d\sigma}+\frac{\partial\phi(R;\rho)}{\partial R}\frac{dR}{d\sigma} \\
&= \frac{\partial\phi(R;\rho)}{\partial\rho}\frac{d\rho}{d\sigma} \\
&= A\left(\frac{(\rho-1)+(\nu-1)}{\rho+\nu}\right)^{\nu}2^{\frac{\nu-1}{\rho-1}}\frac{1}{(\rho-1)+(\nu-1)} \\
&\quad\times\left(\frac{\nu-\rho}{((\rho-1)+(\nu-1))(\rho+\nu)}-(\nu-1)\frac{ln(2)}{(\rho-1)^2}\right)\frac{\rho^2}{\nu}
\end{aligned}
$$

Evaluate at $\sigma=0$:

$$
\begin{aligned}
\frac{d\kappa(0)}{d\sigma}|_{\sigma=0} &= -A\left(\frac{\nu-1}{\nu}\right)^{\nu-1}\frac{ln(2)}{\nu-1} \\
&< 0
\end{aligned}
$$

At $\sigma\in(0,1)$, the derivative is clearly negative.

At the other endpoint of $\sigma=1$, when I take the limit as $\rho\uparrow\infty$:

$$\frac{d\kappa(0)}{d\sigma}|_{\sigma=1} = -\frac{A}{\nu}.$$

So finally, the result obtains:

$$\frac{d\kappa(0)}{d\sigma} < 0.$$

1.9.9 Appendix 9: Properties of the Strategic Substitutes Effect

At the endpoint of $\sigma = 0$, the strategic substitutes effect is equal to 0:

$$\begin{aligned}
\{(\pi(1,1)-\pi(0,1))-(\pi(1,0)-\pi(0,0))\}|_{\sigma=0} &= (\pi(0,0)-\pi(0,1))|_{\sigma=0} \\
&\quad -(\pi(1,0)-\pi(1,1))|_{\sigma=0} \\
&= (\kappa(0)-\kappa(-1))|_{\rho=\nu} \\
&\quad -\left(\kappa(1)\gamma^{\nu-1}-\kappa(0)\gamma^{\nu-1}\right)|_{\rho=\nu} \\
&= \frac{A}{\nu}\left(\frac{\nu-1}{\nu}\right)^{\nu-1}\left(1-1-\left(\gamma^{\nu-1}-\gamma^{\nu-1}\right)\right) \\
&= 0.
\end{aligned}$$

At the endpoint of $\sigma = 1$, the strategic substitutes effect is negative:

$$\begin{aligned}
\{(\pi(1,1)-\pi(0,1))-(\pi(1,0)-\pi(0,0))\}|_{\sigma=1} &= \lim_{\rho\uparrow\infty}\left(\kappa(0)\gamma^{\nu-1}-\kappa(-1)\right) \\
&\quad -\lim_{\rho\uparrow\infty}\left(\kappa(1)\gamma^{\nu-1}-\kappa(0)\right) \\
&= 0\cdot\gamma^{\nu-1}-0-\left(A\left(\frac{\gamma-1}{\gamma}\right)-0\right) \\
&< 0
\end{aligned}$$

CHAPTER 2

Knowledge Transfers and Market Overlap: Technology Licensing and Inventor Job Transitions

2.1 Introduction

In this paper I explore the relationship of knowledge transfer and product substitutability in two channels of knowledge diffusion: technology licensing and inventor job mobility. Both technology licensing deals and inventor job transitions can be classified as voluntary transfers in a market for knowledge between firms. Whereas technology licensing is a straightforward example of a bilateral, compensated transfer of knowledge, even for inventor mobility there is both theory and empirical evidence to support its treatment as a bilateral sale of knowledge. Møen (2005) provides empirical evidence of the use of compensation contracts that allow firms to internalize the returns to their R&D efforts by underpaying workers upfront, before they learn useful knowledge and change jobs.

Understanding the empirical relationship of market-mediated knowledge transfers and product substitutability may aid in determining the relative importance of proposed determinants of knowledge transfer. For example, Kim and Vonortas (2006b) argue that matching frictions take precedence over strategic product market concerns in determining the incidence of technology licensing. Within an Industrial Organization-centric approach, Arora and Fosfuri (2003) assume that entrant-licensees will necessarily occupy the same product-technology niche as the licensor, such that competitive and technological determinants of knowledge transfer are not separately characterized. With regards to inventor mobility, there is a basic question of how effective knowledge transfer via R&D worker mobility is, and whether mobility-generated spillovers are internalized. Finally, any analysis of the unconditional frequency of knowledge transfer between firms will implicitly reflect the incidence of innovation that creates knowledge in the first place.

I use cross-industry data on technology licensing and inventor job mobility to show that the incidence of knowledge transfer in both samples follows a rising-then-falling pattern in product market overlap. Market overlap serves as the empirical analogue to substitutability in this cross-industry comparison of firms. This non-monotonic relationship is obtained when controlling for technological overlap, which displays considerable independent variation with respect to product market overlap. Technology licensing deals are taken from SDC Platinum, Joint Ventures and Alliances component, and span the years 1985-2004. Inventor job transitions are inferred from a USPTO patent dataset in which individual inventors have been identified across their patents by the Fung Institute at Berkeley using a clustering algorithm. This spans 1982-2001. The overlap measures are taken from Bloom et al. (2013) and are calculated using data pooled across years 1980-2004;[1] the market overlap measure compares the bilateral firm similarity in how firms' total sales are distributed across 4-digit SIC industries, and the technological overlap measure compares similarity in the distribution of firms' patents across USPTO patent classes.

In Chapter 1 of this dissertation I show that a rising-then-falling empirical pattern is consistent with a environment in which firms strategically adjust both their innovation and knowledge transfer as a function of product market substitutability. If product market overlap is interpreted in a reduced-form manner as substitutability, then the falling portion of the relationship is due to a lack of gains from trade in knowledge between firms with highly substitutable products. The rising portion of the relationship will occur if greater product substitutability between competitors spurs more innovation and creates more opportunities for knowledge transfer.

With respect to the determinants of licensing that are proposed in the literature, these novel empirical results demonstrate that strategic considerations are of prime importance in governing knowledge flows. In particular, an emphasis on matching frictions alone would yield an incidence of knowledge transfer that rises in product market overlap. Such a rising relationship

[1]The inventor transitions dataset is merged against proximity measures calculated 1980-2001. This better corresponds to the timespan of the inventor transitions.

is what Kim and Vonortas (2006b) find using the same SDC licensing dataset with controls for market and technological overlap. However, they do not use a continuous measure of product market overlap and do not look for a non-monotonic relationship. This paper demonstrates that the relationship of knowledge transfer to product overlap is more accurately characterized as rising-then-falling. Fosfuri (2006) conducts an empirical analysis on licensing within the chemical industry, in which product-technology niches coincide. As such, their scope of investigation does not include licensing across industries and therefore does not isolate the effect of strategic competitive concerns on licensing. Likewise, Kim and Vonortas (2006a) assume that product-technology niches coincide in their analysis of out-licensing behavior of licensors using the SDC dataset. Finally, no other paper in the empirical literature has addressed the incidence of innovation which is implicitly reflected in the knowledge transfer data.

Lastly, the qualitative results of this paper constitute evidence that inventor transitions are governed by the same strategic considerations that apply to licensing deals, suggesting both the importance of such mobility-based knowledge flows and the use of compensation schemes for controlling and internalizing spillovers. As such, it contributes to an extensive literature documenting the knowledge and productivity diffusion that occurs through job mobility of skilled workers. Almeida and Kogut (1999) establish a link between inventor mobility and patent citations of prior art. Stoyanov and Zubanov (2012) show that hiring workers from more productive source firms raises the destination firms' productivity. Kaiser et al. (2015) find that R&D labor mobility increases the total patenting of source and destination firms. As previously mentioned, Møen (2005) finds evidence of both transferable learning and spillover internalization in the phenomenon of workers at R&D intensive firms being underpaid relative to their outside option.

The paper is organized in the following fashion. Section 2 describes the product market and technological overlap measures. Section 3 presents the data sources for licensing and inventor transitions, as well as a model of labor spillover-internalization. Section 4 explores the relationship between market and technological overlap in the sample. Section 5 details the sample construction and baseline results for a probit regression that explains the incidence of knowledge transfer. Section 6 revisits sample construction and presents a preferred sample that excludes pairs of firms with low technological overlap. Section 7 compares linear through quartic polynomial specifications using the preferred sample, and demonstrates the robustness of a non-monotonic pattern of knowledge transfer. Section 8 concludes.

2.2 Product Market and Technological Overlap

The empirical analogue of product substitutability is product market overlap. The first step toward constructing a bilateral measure of market overlap is to compute a vector S_i that describes how the cumulative, price-deflated sales of firm i are distributed across SIC 4-digit level industries. The elements of S_i are expressed in shares of firm i's total sales. I collect the underlying sales data at a parent company level from Compustat Business Line Segments between 1980-2004.

As in Bloom et al. (2013), for firms i and j I use vectors S_i and S_j to calculate a bilateral correlation of the pattern of sales distribution across industries. However, I employ their Mahalanobis measure, which uses a matrix Ω^{SIC} to account for substitutability across pairs of industries.[2] The introduction of this matrix Ω^{SIC} means that the Mahalanobis measure can take on a maximum value greater than 1. The resulting product market overlap between firms i and j is given by SIC_{ij}^{M}, defined below:

$$SIC_{ij}^{M} = \frac{S_i \Omega^{SIC} S_j'}{\left(S_i S_i'\right)^{\frac{1}{2}} \left(S_j S_j'\right)^{\frac{1}{2}}} \geq 0.$$

A larger value of SIC_{ij}^{M} indicates greater product market overlap.

In order to isolate the effect of product substitutability, an empirical analysis must control for technological compatibility between firms. I use the same Mahalanobis methodology to calculate a $TECH_{ij}^{M}$ measure of bilateral technological overlap. Let T_i be a vector which describes how firm i's cumulative patents from 1980-2004 are distributed across USPC patent

[2] If Z_a is the distribution of industry a's total sales across firms (expressed in shares), then Ω^{SIC} matrix entry a, b is given by $\Omega_{a,b}^{SIC} = \frac{Z_a Z_b'}{\left(Z_a Z_a'\right)^{\frac{1}{2}} \left(Z_b Z_b'\right)^{\frac{1}{2}}}$.

classes. Elements of T_i are expressed in shares of firm i's total patents. $TECH_{ij}^M$ is defined as follows:

$$TECH_{ij}^M = \frac{T_i \Omega^{TECH} T_j'}{\left(T_i T_i'\right)^{\frac{1}{2}} \left(T_j T_j'\right)^{\frac{1}{2}}} \geq 0.$$

A larger value of $TECH_{ij}^M$ corresponds to greater technological overlap.

2.3 Knowledge Transfer Data

2.3.1 Technology Licensing Deals

Data on technology licensing deals comes from SDC Platinum's Alliances/Joint Ventures dataset, which flags alliances with licensing content. Anand and Khanna (2000) were the first to use this dataset in an empirical analysis of licensing, but focused on systematic differences in licensing practices between industries. I have screened deals to exclude retail distribution agreements, trademark or television/movie character licensing, and regulatory telecommunications licenses. Remaining deals consist of technology licenses. After merging with the product market and technological overlap measures along with other pair-specific covariates, the sample contains 1,756 bilateral deals with completed/signed status between 1985-2004.[3] These deals occur between 1,603 firm-pairs, which are formed from 1,008 unique firms.

2.3.2 Inventor Job Transitions: Theory

Before presenting the data and inference, I demonstrate that inventor job transitions, which represent a trilateral bargaining problem, can in fact be reduced down to a bilateral bargaining problem between source and destination firm in which the source firm is able to internalize the knowledge spillover. I develop a simple model that is based on the two-stage duopoly game in Chapter 1. To simplify the exposition, assume that the innovation cost c falls in an intermediate

[3] All of the bilateral licensor-licensee relationships present in licenses that had multiple licensees were represented as their own bilateral deals.

range such that firms play a symmetric mixed strategy equilibrium. The timing is as follows:

1. In the first stage, each firm begins at a normalized technology level $H_i = 0$. It hires a skilled worker both to deploy the firm's technology in production and to participate in any R&D that would improve the firm's technology. The firms choose symmetric probability p of innovating at fixed cost c. They also set a wage $w(b, p)$ to pay their worker who has outside option b.

2. In the second stage, if one firm innovated and their rival did not, then the rival contacts the innovator's worker to poach them. Firms bid a bonus Λ_i to be paid to the worker, and the firm that bids the highest bonus Λ^* wins. Successful poaching means that both firms now have the technology.

Afterward, firms make profits $\pi(H_i, H_{-i}; \sigma)$ in the product market that are a function of final technology levels $H_i \in \{0, 1\}$ and substitutability $\sigma \in [0, 1]$. Both firms and workers are risk-neutral.

In the second stage, the valuation of the innovating firm for keeping the worker is $\pi(1,0) - \pi(1,1)$. Likewise, the valuation of the poaching firm for obtaining the worker is $\pi(1,1) - \pi(0,1)$. These valuations will also be a function of σ.

First, suppose that the poaching firm's valuation is higher - σ determines whether the innovator or rival have the higher valuation. In equilibrium, the innovating firm will bid up to $\pi(1,0) - \pi(1,1)$ and establish a floor on the winning bonus Λ^*. The poaching firm's bid will depend on the worker's bargaining power $\theta \in [0, 1]$. If $\theta = 0$, then the poaching firm pays bonus $\Lambda^*(\theta, \sigma) = \pi(1,0) - \pi(1,1)$. At the other extreme, if $\theta = 1$ then the poaching firm pays their full valuation $\Lambda^*(\theta, \sigma) = \pi(1,1) - \pi(0,1)$. If surplus S is defined as $S = \pi(1,1) - \pi(0,1) - (\pi(1,0) - \pi(1,1))$, then the bonus can be written as $\Lambda^*(\theta, \sigma) = \pi(1,0) - \pi(1,1) + \theta S$.

Firms in the first stage will anticipate that the poaching firm will win the worker and pay the bonus $\Lambda^*(\theta, \sigma)$ in the second stage. Therefore in the first stage, firms will pay the worker $w = b - p(1-p)\Lambda^*$, where the amount of underpayment $p(1-p)\Lambda^*$ corresponds to the worker's expected poaching bonus. A firm expects a net gain of $p(1-p)(\Lambda^* - (\pi(1,0) - \pi(1,1))) =$

$p(1-p)\,\theta S$ from the combination of being the sole innovator, having their worker poached, and underpaying that worker in advance. This is exactly the gain expected by a firm from directly licensing a technology to its rival, if the licensor has bargaining power θ. Likewise, the bonus paid by the poaching rival is equal to the lump-sum fee paid by a licensee. Thus this setup is equivalent, in expectation, to a bilaterally bargained licensing deal between the innovating and poaching firms and leads to the same choice of p and incidence of knowledge transfer q as in Chapter 1.

Alternatively, if the innovating firm's valuation is higher, then the innovating firm outbids the rival and keeps the worker in the second stage. In the first stage, the firm pays the worker $w = b - p(1-p)\,\Lambda^*$. This is equivalent, in expectation, to no poaching existing at all. Thus this setup is equivalent, in expectation, to the absence of knowledge transfer that occurs when $S < 0$ in Chapter 1.

This spillover internalization mechanism, which is based on Pakes and Nitzan (1983), shows that inventor job transitions can be mapped into a bilaterally bargained sale of knowledge between source and destination firm. There are limits to such a correspondence, most notably the assumption of risk neutrality, as well as the inability of workers to accept a potentially negative wage w. However, this exercise does demonstrate that firms possess the means to at least partially appropriate innovations even with labor mobility. Møen (2005) provides empirical evidence that R&D intensive firms employ such compensation contracts to force workers to internalize knowledge spillovers.

2.3.3 Inventor Job Transitions: Data

Inventor job transitions between firms must be inferred from USPTO patent data. From patent data alone it is not possible to determine the complete tenures of inventors at firms; instead one can only observe when an inventor patents at one firm and then subsequently patents at another. In line with the literature, I interpret this as a job transition between firms. However, this task of finding transitions is complicated by the fact that inventors are not assigned a unique identifier by the USPTO across patents. For this reason, I use a dataset from the

Fung Institute of Engineering at UC Berkeley that employs a clustering algorithm to identify unique inventors across patents. Their data and clustering algorithm are documented in Fierro et al. (2014). Baslandze (2015) uses a similar dataset from a predecessor patent project at the Harvard Dataverse Network, but uses that data to identify and characterize spinouts.

I impose additional filters and constraints on the inventor transitions data so that it can be interpreted as a knowledge transfer between competing firms. I keep transitions where the time between the last patent at the previous firm and the first patent at the new firm falls between 9 months and 5 years. The lower boundary of 9 months is to guard against faulty identification of inventors in which two distinct individuals are grouped together and are mistaken as one inventor patenting contemporaneously at two different firms. The upper boundary of 5 years only removes the far right tail and is intended to screen out long delays in knowledge transmission that may render the knowledge obsolete. Finally, I remove transitions that are due to mergers and acquisitions or spinoffs. This is because the focus is on poaching between rival firms that continue to compete in the same differentiated product market. Because of the time-intensive task of removing M&A and spinoffs, I limit the inventor transitions sample to 1982-2001. I merge against pair-specific covariates defined over the period 1980-2001 and have 6,006 inventor transitions between 1982-2001. These transitions occur between 3,774 pairs of firms that are formed from 1,568 unique firms.

2.4 Relationship of Market and Technological Overlap

In order for product substitutability to be isolated from technological compatibility as a strategic concern for firms, there must be some degree of independent variation between the market and technology overlap measures.

In the sample of 1,603 technology licensing firm-pairs, the correlation between the product market overlap measure SIC_{ij}^{M} and technological overlap measure $TECH_{ij}^{M}$ is 0.3472. The scatterplot of SIC_{ij}^{M} vs. $TECH_{ij}^{M}$ is in Figure 2.1.

In the sample of 3,774 inventor transition firm-pairs, the correlation between SIC_{ij}^{M} and

$TECH_{ij}^{M}$ is 0.4936. The scatterplot of these measures is depicted in Figure 2.2.

There are two takeaways. First, the correlations between SIC_{ij}^{M} and $TECH_{ij}^{M}$ are appreciably greater than zero, which conforms with intuition. However, a surprising amount of variation exists in product market overlap independent of technological overlap.

As a specific example, Motorola's relationships to IBM and Intel illustrate that pairs of firms can be technologically compatible and yet vary significantly in their market overlap. Motorola is very closely related in technology space to both companies: $TECH_{Moto,IBM}^{M} = 0.98$ and $TECH_{Moto,Intel}^{M} = 1.13$. In product market space, Motorola has almost no overlap with IBM: $SIC_{Moto,IBM}^{M} = 0.01$. On the other hand, Motorola and Intel do overlap in the SIC 3674 "Semiconductors and Related Devices" industry, such that $SIC_{Moto,Intel}^{M} = 0.33$.[4]

More broadly, both samples share the same top five industries as aggregated at the SIC 3-digit level, though not in the same order. For licensing, they are, in descending order, SIC 283 "Drugs," SIC 737 "Computer Programming, Data Processing, And Other Computer Related Services," SIC 357 "Computer and Office Equipment," SIC 367 "Electronic Components and Accessories," and SIC 384 "Surgical, Medical, And Dental Instruments And Supplies." For inventor transitions, they are, in descending order, SIC 283 "Drugs," SIC 367 "Electronic Components," SIC 384 "Surgical, Medical, And Dental Instruments," SIC 357 "Computer and Office Equipment," and SIC 737 "Computer Programming and Data Processing."

2.5 Regression Sample Construction and Probit Regression

The goal in assembling a dataset of technology licensing deals and inventor job transitions is to determine the relationship between the incidence of these examples of knowledge transfer and a continuous measure of product market overlap, while controlling for technological overlap. This raises the following question: which firms are potential parties to knowledge transfer? As a baseline for technology licensing, I form all possible pairs between the 1,008 firms that are represented in my set of licensing deals.[5] This yields 486,389 pairs, of which 1,603 pairs

[4] These Mahalanobis measures were calculated using the technology licensing dataset.

[5] Pairs must have at least one year of overlapping presence in Compustat.

are active - that is, they have a positive number of deals. For inventor transitions I construct a sample of potential parties in the same fashion, yielding 1,115,435 pairs, of which 3,774 are active.

Of the 1,603 active technology licensing pairs, the overwhelming majority, 1,489 pairs, have only a single deal. For inventor job transitions a similar situation exists; out of 3,774 active pairs, 2,905 pairs only have one transition. Given the large number of zeros, and the preponderance of single instances of knowledge transfers among active pairs, a probit regression is used to explain positive vs. zero instances of knowledge transfer in both licensing and job transitions samples.

Let N_{ij}^{LIC} denote the number of licensing agreements between firms i and j in the period 1985-2004. Then I estimate

$$Pr\left(N_{ij}^{LIC} > 0\right) \equiv \Phi\left(X_{ij}'\beta^{LIC}\right),$$

where the vector X_{ij} consists of the following variables:

- SIC_{ij}^{M}: the explanatory variable of interest
- $\left(SIC_{ij}^{M}\right)^2$: a quadratic term to detect non-monotonicity
- $TECH_{ij}^{M}$: a control for technological overlap
- $GEOG_{ij}$: a control for geographic overlap, defined as the correlation of the distribution of two firms' patent-inventor observations across the 50 U.S. states and foreign countries
- $TECH_{ij}^{M} \cdot GEOG_{ij}$: the interaction of technological and geographic overlap
- $Exposure_{ij}$: the number of overlapping years in Compustat
- $Minavgxrd_{ij}$: equal to $min\{\overline{XRD_i}, \overline{XRD_j}\}$ where $\overline{XRD_i}, \overline{XRD_j}$ are firm i and j's average R&D expenditures reported in Compustat during their period of overlap
- A constant term

Likewise, let N_{ij}^{INV} denote the number of inventor transitions between firms i and j in the period 1982-2001. The probit estimation and the covariates are defined analogously to licensing.

Summary statistics for the covariates in both samples are found in Table 2.1. Note that while SIC_{ij}^{M} and $TECH_{ij}^{M}$ can be greater than 1, for practical purposes they vary between 0 and 1.

The baseline probit regression results are reported in Table 2.2 for both licensing and inventor transitions in specifications $1A$ and $1B$, respectively. A significant, rising-then-falling relationship between knowledge transfer and market overlap is revealed here with the positive sign on SIC_{ij}^{M} and negative sign on $\left(SIC_{ij}^{M}\right)^2$ in both samples. The falling part of the relationship is consistent with a lack of gains from trade in knowledge between firms with highly substitutable products.

The rising portion of the relationship can be explained if the intellectual property regime allows buyers of knowledge to capture most of the gains from trade in knowledge. These gains from trade decrease in product substitutability, prompting firms to switch from buying knowledge to performing more of their own innovation. As innovation increases in substitutability, this creates more opportunities for knowledge transfer. A complete theoretical explanation is found in Chapter 1.

The signs on the other coefficients are in line with intuition. In particular, knowledge transfer increases in technological overlap.

2.6 Sample Construction, Revisited

I conduct robustness checks on the non-monotonic relationship found above by imposing stronger requirements on the overlap of potential counterparties. The empirical analysis of knowledge transfer between firms is predicated on a sufficient degree of technological compatibility, such that strategic market competition becomes a first-order consideration in determining the incidence of knowledge transfer. However, there are many firm-pairs, both active and not, that have a very low degree of technological overlap in the sample. In addition, while a thorough analysis of strategic market competition would include pairs of firms with no market overlap, a conservative approach would be to confirm that the large fraction of firm-pairs that are non-overlapping in product space are not driving the results in the two samples.

As a first pass, I eliminate firm-pairs that have zero technological and product market overlap. It is already the case that virtually all firm-pairs in both samples satisfy $TECH_{ij}^{M} > 0$. However, imposing $SIC_{ij}^{M} > 0$ eliminates roughly one-sixth of all firm-pairs in both samples while barely affecting the number of active firm-pairs. Probit regression results are largely unchanged for licensing in specification $2A$ and inventor transitions in specification $2B$.

Beyond eliminating firm-pairs with zero overlap, I also examine whether firm-pairs with positive but low technological compatibility possess a sufficient degree of technological compatibility to treat strategic product market competition as one of their primary considerations in innovation and knowledge transfer. To this end, I split each sample into two subsamples. The first contains firm-pairs with low technological overlap. The second contains the rest of the firm-pairs, which I term the high overlap subsample. I choose the technological overlap cutoff between the low and high technological overlap subsamples as the 25th percentile of $TECH_{ij}^{M}$ among active firm-pairs. For licensing the 25th percentile among active firm-pairs is equal to 0.1913 and for inventor transitions it is equal to 0.1982.

In comparing the subsamples, there are two differences that stand out, the first regarding the unconditional probability of observing knowledge transfer. In the low overlap subsamples, the unconditional probability of any licenses or inventor transitions is 0.1%. In the high overlap subsamples, the corresponding probability is 1.5% for licensing and 2.1% for inventor transitions. The order of magnitude difference indicates that the lack of technological compatibility is, unsurprisingly, a major impediment to knowledge transfer in the low overlap subsamples.

The second notable difference in subsamples concerns the importance of technological overlap in predicting the incidence of knowledge transfer. In Table 2.3, I present results of probit regressions on the split samples. In specifications $3A$ and $3B$, I analyze the low technological overlap subsamples for licensing and inventor transitions, respectively. The magnitude of the coefficient on $TECH_{ij}^{M}$ is very large, greater than 4 in both $3A$ and $3B$. In specifications $4A$ and $4B$, I analyze the high overlap subsamples for licensing and inventor transitions samples, respectively. Here the coefficient on $TECH_{ij}^M$ is noticeably lower than it is in earlier regression specifications 1 and 2.

In Figure 2.3, I plot the marginal effect of $TECH_{ij}^M$, as estimated separately in the two subsamples. The marginal effects are computed at the subsample means of the other explanatory variables, while $TECH_{ij}^M$ is allowed to vary.[6] The marked discontinuity at the subsample boundary confirms that the large difference in $TECH_{ij}^M$ probit coefficients between the two subsamples does translate into a substantial difference in the marginal effect.

The relatively large effect of $TECH_{ij}^M$ for technologically distant potential counterparties complements the finding that their unconditional probability of knowledge transfer is an order of magnitude lower. Together these results suggest that for such potential partners, finding a minimal level of technological compatibility is the overriding concern in knowledge transfer. It can be argued that these potential counterparties do not possess a sufficient degree of technological compatibility to warrant a commensurate focus on strategic competition. In the remaining analysis the low technological overlap pairs are dropped and the subsamples underlying specifications $4A$ and $4B$ are adopted as the preferred regression samples. In these preferred samples, the regression results indicate that a statistically significant rising-then-falling pattern of knowledge transfer still exists.

2.7 Robustness of Regression to Polynomial Specification

Another feature of the regression analysis that warrants more careful examination is the fact that active firm-pairs are bunched close to $SIC_{ij}^M = 0$ and $SIC_{ij}^M = 1$ and relatively sparse in the middle of the range in both samples. This is demonstrated in the scatterplots of Figures 2.1 and 2.2. The bunching at extreme values of SIC_{ij}^M may mean that the non-monotonic relationship described above is in fact largely determined at the endpoints of product market overlap. That calls into question the robust existence of a peak in incidence in the middle of the SIC_{ij}^M range.

To investigate this further, I use the preferred samples underlying specifications $4A$ and

[6] The range is from 0 to the 99th percentile of $TECH^M$ in the high overlap subsample.

4B and run a series of regressions with increasing orders of polynomials in SIC_{ij}^{M}, starting with a linear specification and ending with a quartic specification. In both samples, all of the coefficients on SIC_{ij}^{M} polynomial terms are significant at the 1% level up through the second-order term of the quartic specification. In the licensing sample, the third- and fourth-order terms of the quartic specification are significant at the 5% level. In the inventor transitions sample, the third- and fourth-order terms of the quartic specification are not statistically significant. Results for the polynomial specifications are provided in Table 2.4. In Figure 2.4, I plot the probability that is predicted by varying SIC_{ij}^{M} and holding other variables fixed at their means, first for licensing and then for inventor transitions.

The relationship between knowledge transfer and SIC_{ij}^{M} can be best described as non-monotonic: first rising, then falling. In the range $SIC_{ij}^{M} \in [0, 1]$, it is true across quadratic, cubic, and quartic specifications for both samples that the incidence first rises to a peak, and then ends lower than the peak. In the range $SIC_{ij}^{M} \in (1, 1.13)$, the 95% confidence interval for the quartic specification spreads out and largely encompasses the diverging point estimates given by the other polynomial specifications in both samples. There is less precision in this region because there are relatively few active and potential pairs. In particular, for the licensing sample there are no active pairs toward the end of this $SIC_{ij}^{M} \in (1, 1.13)$ interval, as indicated by the dark grey vertical band. The inconclusiveness in this far-right interval does not appear to contradict the clear rising-then-falling pattern in $SIC_{ij}^{M} \in [0, 1]$.

In evaluating the economic significance of the non-monotonicity, the conservative approach would be to take the range of SIC_{ij}^{M} to be $[0, 1]$. Using the cubic specification in Figure 2.4 as a point of departure for inventor transitions, the probability of observing a positive number of transitions roughly doubles from either endpoint to the peak. The absolute gain is one percentage point. The size of the effect is smaller for licensing, but is nevertheless of the same order of magnitude.

Another, more persuasive way of gauging the economic significance of strategic considerations is to return to regression specifications 4A and 4B. The probit coefficients on the linear and quadratic SIC_{ij}^{M} terms are of the same order of magnitude as the coefficient on the

$TECH^M_{ij}$ term, and both the market and technological overlap measures are defined on the same scale. This comparison demonstrates that strategic competition is quantitatively as important as technological compatibility in determining the pattern of knowledge transfer between firms.

2.8 Conclusion

In this paper I combine measures of voluntary knowledge transfer with market and technological overlap measures to show that inter-firm knowledge transfer is shaped as much by strategic competitive concerns as by basic considerations of technological compatibility. This paper is the first to find that the relationship of knowledge transfer to market overlap is non-monotonic. The rising-then-falling relationship suggests that analyses of knowledge transfer must account for the innovation that creates transferable knowledge in the first place.

The qualitatively similar patterns of transfer for licensing and inventor transitions are evidence for the importance of inventor transitions as a channel of knowledge transfer. The similarity also lends weight to existing evidence in the literature that firms can use compensation schemes to appropriate gains from innovation even when R&D workers can leave for another firm.

2.9 Tables and Figures

Figure 2.1: Scatterplot of SIC_{ij}^{M} vs. $TECH_{ij}^{M}$ for Licensing

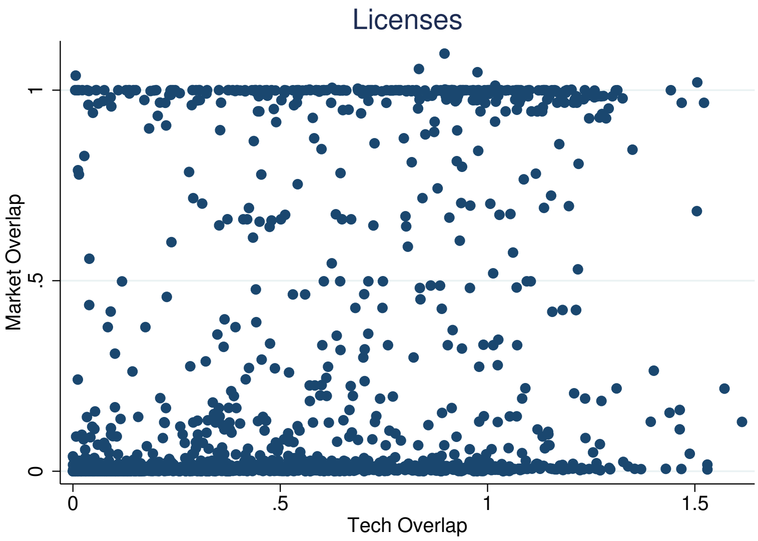

Figure 2.2: Scatterplot of SIC_{ij}^{M} vs. $TECH_{ij}^{M}$ for Inventor Transitions

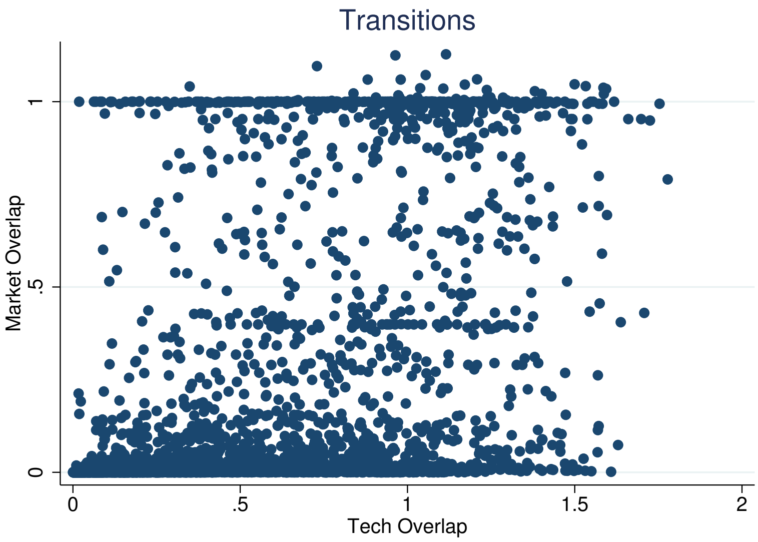

Figure 2.3: Marginal Effect of $TECH_{ij}^{M}$

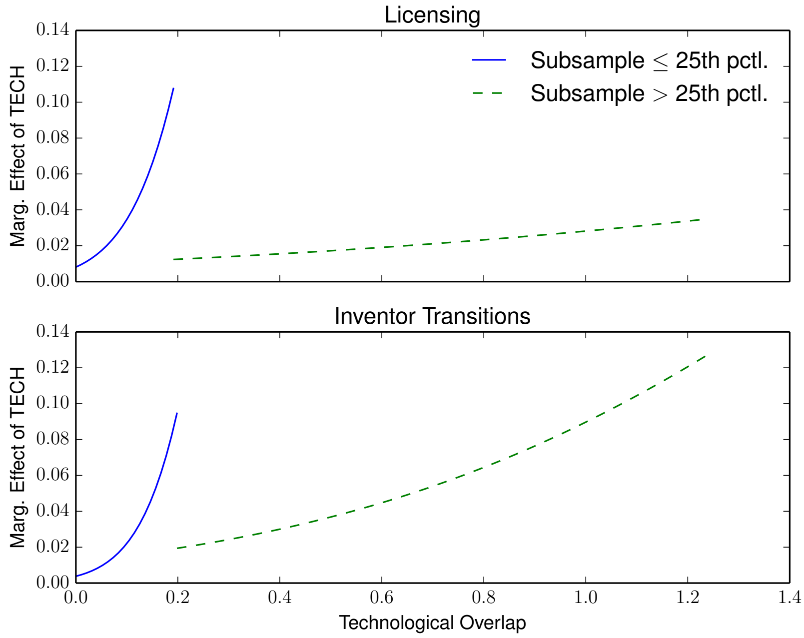

Figure 2.4: Predicted Probabilities of Licensing and Inventor Transitions

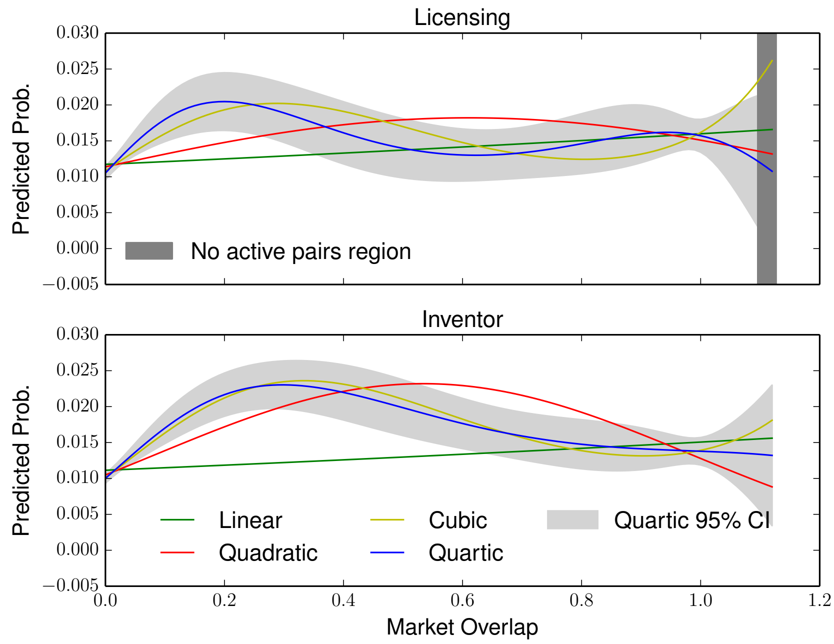

Table 2.1: Summary Statistics for Covariates

Variable	Sample	Min	1st pctl.	Mean	99th pctl.	Max	Stdev.
SIC^{M}_{ij}	licenses	0	0	0.0544	1	1.1268	0.2084
	transitions	0	0	0.0253	0.9998	1.1279	0.1373
$TECH^{M}_{ij}$	licenses	9.75E-10	4.55E-5	0.1134	1.0663	1.9469	0.2283
	transitions	0	2.116E-4	0.0919	0.9677	1.9766	0.1794
$GEOG_{ij}$	licenses	0	0	0.1590	0.9991	1	0.2974
	transitions	0	0	0.1463	0.9981	1	0.2765
$Exposure_{ij}$	licenses	1	1	9.7182	20	20	5.0322
	transitions	1	1	8.9968	20	20	5.5686
$Minavgxrd_{ij}$	licenses	0	0	44.8259	872.37	7228.656	234.772
	transitions	0	0.0488	37.6650	663.7384	8333.333	190.0371

Table 2.2: Regression Results for Baseline and Restricted Positive $TECH^{M}_{ij}$, SIC^{M}_{ij} Specifications

	$(1A)$	$(1B)$	$(2A)$	$(2B)$
	$N^{LIC}_{ij} > 0$	$N^{INV}_{ij} > 0$	$N^{LIC}_{ij} > 0$	$N^{INV}_{ij} > 0$
SIC^{M}_{ij}	1.197***	1.431***	1.141***	1.363***
	(0.168)	(0.125)	(0.168)	(0.126)
$\left(SIC^{M}_{ij}\right)^2$	-1.067***	-1.416***	-1.016***	-1.350***
	(0.170)	(0.129)	(0.170)	(0.129)
$TECH^{M}_{ij}$	1.061***	1.468***	1.024***	1.444***
	(0.0289)	(0.0232)	(0.0292)	(0.0236)
$GEOG_{ij}$	0.111***	0.625***	0.104**	0.624***
	(0.0404)	(0.0249)	(0.0416)	(0.0259)
$TECH^{M}_{ij} \cdot GEOG_{ij}$	0.0822	0.0825*	0.0858	0.0728
	(0.0679)	(0.0450)	(0.0692)	(0.0461)
$Exposure_{ij}$	0.0284***	0.0501***	0.0280***	0.0498***
	(0.00175)	(0.00113)	(0.00178)	(0.00115)
$Minavgxrd_{ij}$	0.000209***	0.000384***	0.000212***	0.000387***
	(1.79e-05)	(1.11e-05)	(1.82e-05)	(1.14e-05)
$constant$	-3.371***	-3.819***	-3.330***	-3.787***
	(0.0254)	(0.0200)	(0.0260)	(0.0207)
McFadden's R^2	0.1465	0.2589	0.1387	0.2513
Restriction on SIC^{M}_{ij}	None	None	> 0	> 0
Restriction on $TECH^{M}_{ij}$	None	None	> 0	> 0
N	486,389	1,115,435	418,667	926,497
$N > 0$	1,603	3,774	1,580	3,694

Table 2.3: Regression Results for Technological Overlap Subsamples

	$(3A)$	$(3B)$	$(4A)$	$(4B)$
	$N_{ij}^{LIC} > 0$	$N_{ij}^{INV} > 0$	$N_{ij}^{LIC} > 0$	$N_{ij}^{INV} > 0$
SIC_{ij}^{M}	1.931***	0.817**	0.606***	1.185***
	(0.387)	(0.364)	(0.185)	(0.133)
$\left(SIC_{ij}^{M}\right)^2$	-1.653***	-0.698*	-0.496***	-1.110***
	(0.408)	(0.396)	(0.187)	(0.136)
$TECH_{ij}^{M}$	4.545***	5.438***	0.468***	0.913***
	(0.303)	(0.231)	(0.0457)	(0.0357)
$GEOG_{ij}$	-0.109	0.651***	0.142*	0.609***
	(0.0921)	(0.0516)	(0.0810)	(0.0521)
$TECH_{ij}^{M} \cdot GEOG_{ij}$	1.292	-0.548	0.0412	0.0855
	(0.892)	(0.498)	(0.113)	(0.0734)
$Exposure_{ij}$	0.0134***	0.0396***	0.0345***	0.0516***
	(0.00308)	(0.00193)	(0.00227)	(0.00147)
$Minavgxrd_{ij}$	0.000125***	0.000292***	0.000267***	0.000447***
	(3.50e-05)	(1.92e-05)	(2.43e-05)	(1.55e-05)
$constant$	-3.474***	-4.037***	-2.922***	-3.394***
	(0.0436)	(0.0367)	(0.0416)	(0.0323)
McFadden's R^2	0.0716	0.1424	0.0521	0.1514
Restriction on SIC_{ij}^{M}	> 0	> 0	> 0	> 0
Restriction on $TECH_{ij}^{M}$	≤ 25th pctl.	≤ 25th pctl.	>25th pctl.	>25th pctl.
N	340,699	791,797	77,968	134,700
$N > 0$	381	885	1,199	2,809

Table 2.4: Regression Results with Polynomial SIC_{ij}^{M} Terms

	Specification 5							
	Licensing $N_{ij}^{LIC} > 0$				Inventor Transitions $N_{ij}^{INV} > 0$			
SIC_{ij}^{M}	0.121***	0.606***	1.947***	3.163***	0.117***	1.185***	2.313***	2.728***
$\left(SIC_{ij}^{M}\right)^2$		-0.496***	-4.595***	-12.14***		-1.110***	-4.745***	-7.308***
$\left(SIC_{ij}^{M}\right)^3$			2.803***	15.88**			2.555***	7.012
$\left(SIC_{ij}^{M}\right)^4$				-6.742**				-2.309
Other covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
McFadden's R^2	0.0515	0.0521	0.0531	0.0534	0.1490	0.1514	0.1521	0.1521
N	77,968	77,968	77,968	77,968	134,700	134,700	134,700	134,700
$N > 0$	1,199	1,199	1,199	1,199	2,809	2,809	2,809	2,809

2.10 Appendix: Marginal Effects

In splitting the licensing and inventor transitions samples into low and high technological proximity subsamples, I calculate the marginal effect of $TECH_{ij}^{M}$ in subsamples as follows:

$$\begin{aligned}\frac{\partial P\left(N_{ij}>0\right)}{\partial TECH_{ij}^{M}} &= \phi(\hat{\beta}_0+\hat{\beta}_1\overline{SIC_{ij}^{M}}+\hat{\beta}_2\overline{SIC_{ij}^{M}}^{2}+\hat{\beta}_3 TECH_{ij}^{M}+\hat{\beta}_4\overline{GEOG_{ij}} \\ &\quad +\hat{\beta}_5 TECH_{ij}^{M}\cdot\overline{GEOG_{ij}}+\hat{\beta}_6\overline{Exposure_{ij}}+\hat{\beta}_7\overline{Minavgxrd_{ij}}) \\ &\quad \times\left(\hat{\beta}_3+\hat{\beta}_5\overline{GEOG_{ij}}\right)\end{aligned}$$

That is, I calculate at the averages of the other explanatory variables within each subsample, but allow $TECH_{ij}^{M}$ to vary.

CHAPTER 3

Voluntary Knowledge Transfers, Competition, and Growth

3.1 Introduction

What determines firms' incentives to innovate and transfer knowledge to other firms? Some channels of knowledge diffusion represent voluntary transfers - technology licensing being one such example - and therefore the incidence of transfer will depend on product market competition. I ask how changes in intellectual property policies affect voluntary knowledge transfer, innovation, and growth across different degrees of product substitutability.

I develop a dynamic duopoly model of sequential innovation that allows endogenous transitions among any number of technology level gaps $\tau \in \mathbb{N}^0$ between duopolists. A firm that is behind the duopoly frontier will be termed a follower, and has two options for improving its knowledge capital level. The first is to innovate at a weakly higher cost than its rival at the frontier, which will be termed the leader. The second is to initiate stochastic imitation contact at a cost, which if successful leads to Nash bargaining and transfer of the frontier knowledge. This contact friction acknowledges the follower's role in knowledge transfer and places limits on the effectiveness of raising the innovator bargaining power to increase innovation.

I use the dynamic model for two purposes, first to explain a novel empirical pattern of knowledge transfer found in Chapter 2 of this dissertation, and second to perform policy counterfactuals on the intellectual property regime. With regard to the first objective, I demonstrate that the model can produce the empirical rising-then-falling relationship between knowledge transfer and product substitutability that is found in Chapter 2. Having shown that the model has empirical relevance, I then use it to evaluate the effect of intellectual property policy changes on innovation, the endogenous technology gap between firms, and the overall impact on the average output growth rate. Policies can be modeled by changing the leader's bargaining power or the follower's innovation and imitation cost parameters, which correspond to patent protection strength, patent breadth, and frictions in licensing or job mobility, respectively.

In allowing the technology gap between firms to be endogenously determined, the dynamic model yields insights on the evolution of industry structure and the role that structure plays in the outcomes of policy counterfactual exercises. At high degrees of substitutability the two firms drift apart in technology levels, because there are no gains from trade in knowledge and therefore transfer does not occur. In the long-run, one of the firms will pull so far ahead that it monopolizes the industry, despite the ability of the follower to innovate. This outcome is characterized by a low level of innovation and a low output growth rate.

At lower degrees of substitutability, raising bargaining power above zero increases leader innovation, but the leader cannot sell their technology unless the follower invests in overcoming the knowledge-transfer friction. This limits the innovation-incentivizing effects of raising the leader's bargaining power. Furthermore, the combination of more leader innovation and less conditional knowledge transfer moves duopolists away from technological equality at the frontier. This shift in the industry structure is anti-competitive with regard to innovation, and it reverses a substantial portion of the gains in the output growth rate that come from increased leader innovation.[1] Policies that lower the follower's cost of innovation or the knowledge transfer friction can increase growth at or immediately below the monopolization boundary, respectively.[2] At lower degrees of substitutability, these follower cost policies result in decreased growth.

This paper is closely related to a literature that analyzes the effects of knowledge transfer on measures of social welfare in a strategic context. Bloom et al. (2013) empirically distinguishes between the product market rivalry and technological spillover effects of innovation and finds that technology spillovers dominate, thus making the case for more R&D. Aghion et al. (2001) and Aghion et al. (2005) model a dynamic game of innovation and spillovers, finding some evidence for a positive role for spillovers in output growth and innovation, respectively. These

[1] The net impact on growth is actually negative at levels of substitutability in the neighborhood of the monopolization boundary.

[2] The follower is assumed to innovate at a weakly higher cost than the leader.

examples in the literature are representative in that they do not take a stand on how knowledge is transferred and assume the diffusion to be a spillover. Acemoglu and Akcigit (2012) build on Aghion et al. (2001) and analyze optimal state-dependent intellectual property protection. They address compulsory licensing, and find that voluntary licensing would never occur in their model environment. In contrast to these examples from the literature, my paper focuses on channels of diffusion that are voluntary and market-mediated and is able to explain qualitative empirical patterns of voluntary knowledge transfer. In that last regard, my paper is set apart from Bessen and Maskin (2009). They evaluate the conditions under which voluntary licensing would occur in a model of sequential innovation, but do not predict or explain an empirical relationship between the incidence of knowledge transfer and product substitutability.

The paper is organized as follows. In Section 2 the dynamic game environment is specified. In Section 3 I present the mechanisms that can create a rising-then-falling relationship of knowledge transfer to substitutability. In Section 4 the calibration of the model is discussed. Section 5 shows that at zero innovator bargaining power, a non-monotonic pattern of knowledge transfer obtains, for the same reasons as discussed in Section 3. Section 6 conducts counterfactual experiments with different aspects of the intellectual property regime. Section 7 concludes.

3.2 Dynamic Model Environment

3.2.1 Demand, Profits, and State Variables

Duopolists engage in price competition and face nested CES demand, where ν is the outer industry elasticity of substitution and $\rho \geq \nu$ is the inner elasticity of substitution between firms j and $-j$. Competition against an implicit outside sector allows for the existence of a positive surplus from knowledge transfer within the duopoly. Instead of varying inner CES parameter $\frac{\rho-1}{\rho} \in \left[\frac{\nu-1}{\nu}, 1\right]$, I define a rescaled substitutability index σ that moves between 0 and 1 as $\frac{\rho-1}{\rho}$ varies:

$$\sigma = \frac{\frac{\rho-1}{\rho} - \frac{\nu-1}{\nu}}{1 - \frac{\nu-1}{\nu}}.$$

Production occurs according to constant marginal cost $\frac{\psi}{H_i}$, where H_i is the knowledge capital of firm i and ψ is normalized to one. The dynamic, continuous-time discount rate on profits is given by r.

I prove in Appendix Section 3.9.1 that the static indirect profit function for firm j, $\pi\left(H_j, H_{-j}\right)$, can be multiplicatively separated into a relative component $\pi_i(\tau)$ and an absolute component $H^{\nu-1}$, where H is the knowledge capital level of the leader, τ is the number of innovations separating the leader from the follower, and subscripts L, F, E refer to Leader, Follower, and Equal, respectively:

$$\pi\left(H_j, H_{-j}\right) = \pi_i(\tau) H^{\nu-1}, i \in \{L, F, E\}.$$

In what follows, the state variables of this dynamic game are given by (H, τ).

3.2.2 Innovation and Knowledge Transfer

Firms have two different means for improving their knowledge capital, the first being innovation. The arrival of an innovation improves a firm's knowledge capital by a multiplicative factor $\gamma > 1$. Innovation may be undertaken by a leader or a follower. The second is imitation via bargained knowledge transfer, which may only be undertaken by a follower. Knowledge transfer brings the follower up to the frontier such that $\tau = 0$.

In this continuous time model, firms choose the Poisson arrival rate of their innovations, subject to a cost that is quadratic in the arrival rate. The cost also scales as a function of the knowledge capital level of the firm, in such a way that a firm's innovation costs and profits grow proportionally as its knowledge capital grows. In particular, a firm at the frontier (either a leader, or neck-and-neck with its rival) chooses an innovation Poisson arrival rate $p_i(\tau)$ at quadratic cost $H^{\nu-1} c \frac{p_i(\tau)^2}{2}$. A follower τ steps behind the frontier chooses innovation arrival rate $p_F(\tau)$ at cost $\left(\frac{H}{\gamma^{\tau}}\right)^{\nu-1} c_F \frac{p_F(\tau)^2}{2}$. Note that a firm at the frontier can innovate with cost parameter c, whereas a follower firm behind the frontier innovates with cost parameter $c_F \geq c$. The follower faces a higher innovation cost because they must innovate around the intellectual property of the leader.

Bargained knowledge transfer is subject to an imitation contact friction that reflects search and matching efforts. A follower chooses the Poisson arrival rate of imitation contact $f(\tau)$ at cost $H^{\nu-1}\eta\frac{f(\tau)^2}{2}$. The cost of imitation contact scales as a function of $H^{\nu-1}$ to keep up with the joint surplus created by knowledge transfer, since surplus will also be a function of $H^{\nu-1}$. When imitation contact arrives, the imitator and innovator Nash bargain over the lump sum payment for knowledge transfer. The innovator's bargaining power is given by θ, and a bargain is struck if and only if the surplus S created by knowledge transfer is positive. A bargaining power of $\theta = 0$ leaves the innovator indifferent to transferring the knowledge to their competitor. On the other hand, $\theta = 1$ leaves the imitator indifferent to receiving the knowledge.

3.2.3 Value Functions

The value function $V_E(H)$ for firms that are equal at the frontier is given by:

$$\begin{aligned} rV_E(H) &= \max_{p_j} \pi_E H^{\nu-1} - H^{\nu-1}\frac{cp_j^2}{2} \\ &\quad + p_j\left(V_L(\gamma H, 1) - V_E(H)\right) + p_{-j}\left(V_F(\gamma H, 1) - V_E(H)\right). \end{aligned}$$

The leader's value function $V_L(H, \tau)$ is given by:

$$
\begin{aligned}
rV_L(H,\tau) &= \max_{p_L(\tau)} \pi_L(\tau) H^{\nu-1} - H^{\nu-1}\frac{cp_L(\tau)^2}{2} \\
&\quad + p_L(\tau)\left(V_L(\gamma H,\tau+1) - V_L(H,\tau)\right) \\
&\quad + p_F(\tau)\left(\left(1_{\tau>1}V_L(H,\tau-1) + 1_{\tau=1}V_E(H)\right) - V_L(H,\tau)\right) \\
&\quad + f(\tau)\theta\bar{S}(H,\tau) \\
s.t. & \\
\bar{S}(H,\tau) &= \max\{2V_E(H) - V_L(H,\tau) - V_F(H,\tau), 0\}.
\end{aligned}
$$

The follower's value function $V_F(H,\tau)$ is given by:

$$
\begin{aligned}
rV_F(H,\tau) &= \max_{p_F(\tau), f(\tau)} \pi_F(\tau) H^{\nu-1} - \left(\frac{H}{\gamma^\tau}\right)^{\nu-1}\frac{c_F p_F(\tau)^2}{2} - H^{\nu-1}\frac{\eta f(\tau)^2}{2} \\
&\quad + p_L(\tau)\left(V_F(\gamma H,\tau+1) - V_F(H,\tau)\right) \\
&\quad + p_F(\tau)\left(\left(1_{\tau>1}V_F(H,\tau-1) + 1_{\tau=1}V_E(H)\right) - V_F(H,\tau)\right) \\
&\quad + f(\tau)(1-\theta)\bar{S}(H,\tau) \\
s.t. & \\
\bar{S}(H,\tau) &= \max\{2V_E(H) - V_L(H,\tau) - V_F(H,\tau), 0\}.
\end{aligned}
$$

In Appendix Section 3.9.2 I prove that there exists a value function solution in which the technological frontier component $H^{\nu-1}$ can be factored out of $V_i(H,\tau)$. Such a solution exists due to the multiplicative separability of the relative and absolute components of profit, along with convenient choices for the scaling of innovation and imitation costs. This yields modified value functions v_i that are a function of relative position τ only:

$$
v_i(\tau) = \frac{V_i(H,\tau)}{H^{\nu-1}}.
$$

Suppose innovation increases the frontier knowledge capital to $H' = \gamma H$. This increases

$v_i(\tau)$ by a factor of $\gamma^{\nu-1}$.

The state variable τ can take on any weakly positive integer value. However, I show in Appendix Section 3.9.3 that $\pi_i(\tau)$ converges asymptotically as $\tau \uparrow \infty$ for $i \in \{L, F\}$, and therefore $v_i(\tau)$ converges asymptotically as well. For computational tractability, the state space is defined as $\tau \in \{0, 1, ..., \bar{\tau}\}$ where $\bar{\tau}$ satisfies $v_i(\bar{\tau}) \approx \lim_{\tau\uparrow\infty} v_i(\tau)$ for $i \in \{L, F\}$.

3.3 Pattern of Knowledge Transfer When $\theta = 0$

Denote the average arrival rate of knowledge transfer as q. Empirically, Chapter 2 uses licensing and inventor transitions data to show that knowledge transfer q follows a rising-then-falling pattern in substitutability σ. Theoretically, Chapter 1 uses a two-stage duopoly game to demonstrate that such a non-monotonic pattern can occur if the innovator's bargaining power θ is close to 0. The intuition for the rising part of the pattern is that when $\theta = 0$, a buyer of knowledge reaps the gains from trade in knowledge, but switches toward their own innovation as these gains from trade decrease in substitutability. More innovation creates more opportunities for knowledge transfer, as substitutability increases.

Here, I show that for $\theta = 0$, the core mechanisms of the dynamic model can also generate a rising-then-falling incidence of knowledge transfer as σ increases. Define $\mu(\tau)$ as the invariant distribution frequency of being in state τ. Then the average arrival rate of knowledge transfer q is given by the following expression:

$$q = \sum_{\tau>0} \mu(\tau) f(\tau).$$

This can be rewritten as the product of the total measure μ_A of all asymmetric states $\tau > 0$, and the weighted-average imitation arrival rate $\bar{f}$ across $\tau > 0$:

$$q = \mu_A \bar{f}.$$

Knowledge transfer q will initially rise in σ if opportunities for knowledge transfer μ_A increase in substitutability σ and the conditional arrival rate $\bar{f}$ does not counteract the effect of μ_A on q. However, q will eventually fall to zero in σ because surplus becomes negative at high σ and $\bar{f}$ declines to zero.

Other flows held equal, if firms innovate more in the equal state $\tau = 0$ as σ increases, then they will spend a greater fraction μ_A of their time in unequal states and there will be more opportunities for knowledge transfer as σ increases. This intuition is captured in the following expression of μ_A as a function of equal state innovation $2p_E$:[3]

$$\mu_A \quad = \quad \frac{2p_E - \mu(1)\, p_F(1)}{2p_E + \bar{f}}.$$

It remains to show that $\frac{d}{d\sigma}(2p_E) > 0$ such that opportunities for knowledge transfer μ_A increase in σ. I also need to demonstrate that the conditional arrival rate $\bar{f}$ does not counteract the effect of μ_A on q. Together, these results would demonstrate that q initially rises in σ. To provide transparent intuition, I briefly analyze a restricted dynamic model in which $\tau \in \{0,1\}$ and the only action that occurs in the asymmetric state $\tau = 1$ is an exogenous arrival rate $f(1)$ of bargained imitation. This exogenous arrival rate $f(1)$ has a function analogous to the imitation contact friction cost η in the full dynamic model.

In such a restricted dynamic model, the first order condition for p_E is given by

$$p_E = \frac{\gamma^{\nu-1} v_L(1) - v_E}{c}.$$

The total derivative $\frac{dp_E}{d\sigma}$ is composed of a direct term as well as a feedback term:

[3]The other flows $\mu(1)\, p_F(1)$ and $\bar{f}$ will not counteract the effect of an increase in $2p_E$. The discouragement effect that falling behind has on follower innovation will limit the size and movements of $\mu(1)\, p_F(1)$ relative to $2p_E$. Furthermore, $\bar{f}$ would have to increase substantially in σ to negate a pattern $\frac{d\mu_A}{d\sigma} > 0$, which will not occur because of its dependence on surplus.

$$\frac{dp_E}{d\sigma} = \underbrace{\frac{\partial p_E}{\partial \sigma}}_{\text{Direct}} + \underbrace{\frac{\partial}{\partial p_E}\left(\frac{\gamma^{\nu-1} v_L(1) - v_E}{c}\right)\frac{\partial p_E}{\partial \sigma}}_{\text{Feedback}}$$

I demonstrate below that $\frac{\partial p_E}{\partial \sigma} > 0$. The result $\frac{dp_E}{d\sigma} > 0$ follows if c is sufficiently large to limit the feedback effect of p_E onto itself.

The partial derivative $\frac{\partial p_E}{\partial \sigma}$ can be written thus, where I hold p_E fixed on the right hand side:

$$\frac{\partial p_E}{\partial \sigma} \propto \frac{d}{d\sigma}\left(\gamma^{\nu-1}\pi_L(1) - \pi_E\right) + \frac{\gamma^{\nu-1} p_E}{r}\frac{d}{d\sigma}\left(\pi_L(1) - \pi_F(1)\right) - \frac{\gamma^{\nu-1} p_E f(1)}{r}\frac{\partial S(1)}{\partial \sigma}.$$

I show in the Appendix Section 3.9.4 that all three terms are at least weakly positive so that $\frac{\partial p_E}{\partial \sigma} > 0$.[4] The intuition for the first two terms is that the gains in static profit from innovating and becoming a leader increase in σ.[5] The intuition for the third term involves the weakening of a negative, disincentive effect. Firms in the equal state are disincentivized to innovate because they can reap the surplus from knowledge transfer if they become a follower. However, this disincentive effect is weakened as substitutability increases and surplus decreases: $\frac{\partial S}{\partial \sigma} < 0$. This intuition is also found in Chapter 1.

Thus far, I have used a restricted dynamic model to show that the increase in innovation p_E creates more opportunities for knowledge transfer μ_A as σ increases. I now demonstrate that the conditional arrival rate $\bar{f}$ can be initially increasing in σ. If so, the increase in $\bar{f}$ works together with the increase in knowledge transfer opportunities so that q will initially increase in σ.

In the full dynamic model, the follower's incentive to invest in conditional arrival rate $f(\tau)$ depends on surplus $\bar{S}(\tau)$, so $\bar{f}$ will follow weighted-average surplus (across τ).[6] In the

[4]The first term $\frac{d}{d\sigma}\left(\gamma^{\nu-1}\pi_L - \pi_E\right)$ is, to a first approximation, weakly positive.

[5]In the first term, becoming a leader is compared against the equal state π_E. In the second term the comparison is against slipping further and becoming a follower π_F.

[6]See the first order condition at given τ: $f(\tau) = \frac{(1-\theta)\bar{S}(\tau)}{\eta}$

restricted dynamic model there is no direct analogue to $\bar{f}$, since there is only one asymmetric state $\tau = 1$ and $f(1)$ is exogenously given. However, given the dependence of $\bar{f}$ on weighted-average surplus, $\bar{f}$ from the full dynamic model can be modeled using surplus $S(1)$ in the restricted model.

In what follows, I show that as σ increases and p_E along with it, a firm in the equal state perceives that there is a greater chance of becoming a follower and reaping the large surplus that exists at low σ. This increases their value v_E, which directly increases surplus and thereby incentivizes greater investment in $\bar{f}$ as σ increases.

Movements in v_E will pass on to surplus. In the restricted dynamic model, surplus can be expressed as a function of v_E as follows:

$$\begin{aligned} S(1) &= 2v_E - v_L(1) - v_F(1) \\ &= \frac{2v_E - \frac{\pi_F(1)+\pi_L(1)}{r}}{\left(1 + \frac{f(1)}{r}\right)}. \end{aligned}$$

The Envelope Theorem shows that the opponent's innovation p_{-j} has an effect on v_E as substitutability increases:

$$\frac{dv_E}{d\sigma} = \frac{\partial v_E}{\partial \sigma} + \frac{\partial v_E}{\partial p_{-j}} \frac{dp_{-j}}{d\sigma}$$

As σ increases, p_{-j} does as well: $\frac{dp_{-j}}{d\sigma} > 0$. If a sufficiently large free-riding effect exists, as expressed in $\frac{\partial v_E}{\partial p_{-j}} \frac{dp_{-j}}{d\sigma} \gg 0$, that may be enough to create an increase in v_E: $\frac{dv_E}{d\sigma} > 0$.[7]

In the restricted dynamic model, the effect of the opponent's innovation is given by the change in value from becoming a follower:

[7]Note that $\frac{\partial v_E}{\partial \sigma} < 0$ over the relevant domain.

$$
\begin{aligned}
r\frac{\partial v_E}{\partial p_{-j}} &= \gamma^{\nu-1}v_F(1)-v_E \\
&\propto \left(\gamma^{\nu-1}\pi_F(1)-\pi_E\right)+\frac{\gamma^{\nu-1}p_E}{r}\left(\pi_F(1)-\pi_L(1)\right)+\gamma^{\nu-1}f(1)\,S(1)\left(1+\frac{p_E}{r}\right)+\frac{cp_E^2}{2}.
\end{aligned}
$$

It is possible for $\frac{\partial v_E}{\partial p_{-j}}>0$ to occur because the third term $\gamma^{\nu-1}f(1)\,S(1)\left(1+\frac{p_E}{r}\right)$ is positive and can be large at low σ where surplus is large.[8] The exogenous arrival rate $f(1)$ must also be sufficiently high for this third term to be large enough. The intuition for this third term is as follows: there may be a net gain to a firm in the equal state when its opponent increases innovation, if as a follower the value $\gamma^{\nu-1}f(1)\,S(1)\left(1+\frac{p_E}{r}\right)$ derived from free-riding off surplus is more than the loss in present discounted value of static profits as represented by the first two terms.

This free-riding effect on v_E may be strong enough to increase surplus as σ increases, and is more likely create such a positive relationship at lower values of σ. Rising surplus leads to a rise in $\bar{f}$ as σ increases. Such behavior of $\bar{f}$ is an important component - together with μ_A - in explaining why the arrival rate q of knowledge transfer can initially rise in σ.

3.4 Solution and Calibration

I solve for a Markov Perfect Equilibrium using the method set forth by Pakes and McGuire (1994). I set the discount rate $r=0.04$, and I normalize the market size parameter of $A=1$ for CES industry demand $Q=AP^{-\nu}$.[9] I set $\theta=0$ because it corresponds to a sharp prediction that the dynamic model can generate a rising-then-falling pattern of knowledge transfer with respect to substitutability. Such a qualitative relationship would match the empirical findings in Chapter 2. I calibrate the remaining parameters: the CES industry demand elasticity ν, the innovation growth factor γ, the innovation cost parameters c and c_F, and the imitation contact cost parameter η.

[8]The first two terms are at most weakly negative, and the fourth term is positive.

[9]Insofar as policy functions and associated costs are concerned, only the ratios $\frac{c}{A}$, $\frac{c_F}{A}$, and $\frac{\eta}{A}$ can be identified.

The growth and cost parameters are calibrated to targets computed as averages across sector level moments in BEA data. In particular, the majority of firms in the licensing and inventor transitions samples fall into broad categories of Manufacturing, Information, and "Professional, Scientific, and Technical Services," defined at a NAICS 2-digit level by the BEA. Within these categories the BEA provides annual output and IP investment data at a more disaggregated NAICS 3-digit sector level. For any particular target moment (across time), I first calculate at the sector level between 1987-2004, and then average the results across all sectors within the 3 broad categories mentioned above to arrive at a calibration target.

I calibrate the growth factor γ to match an average annual real output growth rate of 2.6%, and pick η to match the variance of the real output growth rate at 0.25%. Innovation cost parameters c and c_F are chosen to match an average IP investment to nominal output ratio of 3% and variance of IP investment to nominal output ratio of 0.02%. Details on the calculation of these targets in the model are provided in the Appendix Section 3.9.5.

Industry demand elasticity ν is chosen to target a peak in the incidence of knowledge transfer at $\sigma = 0.25$. Such a value is in line with the probit regression results for licensing and inventor transitions samples in Chapter 2.

The resulting parameter values are as follows: the outer CES parameter is $\frac{\nu-1}{\nu} = 0.29$, the growth factor is $\gamma = 1.14$ (which is similar to the baseline value of 1.135 used by Aghion et al. (2001)), the innovation costs are given by $c = 6.0$ and $c_F = 10.5$, and the imitation contact cost is $\eta = 1.5$. These parameter values deliver model moments that are mostly close to the targets in the data, as seen in Table 3.1.

3.5 Results: Decomposition of Incidence of Knowledge Transfer

With a bargaining power of $\theta = 0$, the calibrated model delivers an incidence of knowledge transfer q that is rising-then-falling in substitutability σ. This result is depicted in Figure 3.1, and is consistent with the qualitative relationships found in licensing and inventor transitions data in Chapter 2. In the following section I decompose the model results to demonstrate that the mechanisms in Section 3.3 explain the non-monotonic relationship.

Here I revisit the decomposition of q:

$$q = \mu_A \bar{f}.$$

In decomposing the contributions of μ_A and $\bar{f}$, there are three things to point out. The first is that μ_A increases in σ and is the primary driver of the initial increase in q. The second is that the behavior of $\bar{f}$ is non-monotonic, first increasing and then decreasing in σ. The behavior of $\bar{f}$ follows that of surplus, which is also non-monotonic because of a free-riding effect explained earlier. Third, monopolization occurs for $\sigma \geq 0.41$, as reflected in μ_A reaching a value of 1 and $\bar{f}$ falling to a corresponding value of 0.

The decomposition is depicted in Figure 3.2, first in levels, and then in a log scale in the bottom panel so that multiplicative contributions are easy to assess.

What drives $\frac{d\mu_A}{d\sigma} > 0$? Recall that μ_A can be written as

$$\mu_A = \frac{2p_E - \mu(1)\, p_F(1)}{2p_E + \bar{f}}.$$

The positive relationship $\frac{dp_E}{d\sigma} > 0$ is the primary driver of the result $\frac{d\mu_A}{d\sigma} > 0$. The role of p_E can be seen in Figure 3.3, where μ_A is plotted, along with the numerator $2p_E - \mu(1)\, p_F(1)$, $2p_E$ which is the main force behind the behavior of the numerator, and the inverse of the denominator $\frac{1}{2p_E + \bar{f}}$. Everything is on a log scale.

The steep decline of q between $\sigma = 0.40$ and $\sigma = 0.41$ reflects the onset of monopolization. For monopolization to occur, it must be the case that a follower firm that is arbitrarily far behind the frontier has no prospect of catching up. This condition is fulfilled in part by what happens to surplus and knowledge transfer as σ increases. For $\sigma \geq 0.41$, a follower that is sufficiently far behind the frontier faces negative surplus and cannot take advantage of knowledge transfer.[10] The cessation of knowledge transfer is coupled with a decreasing incentive for the follower to innovate when it is arbitrarily far behind the frontier, as σ increases past 0.40.

[10]Given high enough σ, surplus defined on static profit, denoted static surplus $S^{\pi}(\tau)$, exhibits a rising-then-falling relationship in τ. The dynamic model accentuates this property of $S^{\pi}(\tau)$.

This, together with a positive and non-negligible leader innovation arrival rate, means that the follower falls irretrievably behind. The innovation policy functions $p_L(\tau)$ and $p_F(\tau)$ of the leader and follower, respectively, are depicted for the monopolization region in Figure 3.4 for $\sigma \in \{0.41, 0.71, 1.0\}$.

3.6 Intellectual Property Policy Counterfactuals

The calibrated model provides suggestive evidence on the value of parameters that are of interest for intellectual property protection. In Figure 3.6 I show that at values of $\theta \geq 0.25$, the calibrated dynamic model does not produce a non-monotonic relationship of q to σ. Thus the bargaining power θ of innovators is likely to be low in practice; this serves as an index of appropriability and is a proxy for the strength of patent protection. The follower innovation cost parameter c_F is 75 percent higher than that faced by firms at the frontier; this can be interpreted as a measure of how difficult it is to innovate around a patent and stands in for breadth of patent protection.

Lastly, the imitation contact cost η is non-negligible. In a survey by Radauer and Dudenbostel (2013), firms report significant difficulties in finding suitable licensing partners, and they make use of informal networks for finding such partners. Alternatively, η may have a labor market interpretation. In Chapter 2 I show that inventor job transitions fit into a paradigm of a bilateral sale of knowledge between firms, and this claim is borne out by the qualitative similarity in the empirical patterns of licensing and inventor transitions. In a labor mobility context, η represents frictions in recruiting R&D workers from rivals in the presence of non-compete agreements and trade secret laws. There is an extensive empirical literature documenting the negative effects that non-compete agreements have on the labor mobility of technical and managerial workers - see Belenzon and Schankerman (2013), Fallick et al. (2006), Garmaise (2009), and Marx et al. (2009) as examples.

The importance of the degree of patent protection strength, patent breadth, and knowledge transfer frictions can be measured through their effects on the average growth rate g of duopoly output. In the Appendix Section 3.9.5 I show that g is the inner product of two vectors: the rate of frontier innovation and the invariant distribution:

$$g \approx \left(\mu(0)\,2p_E + \sum_{\tau=1}^{\tau=\bar{\tau}} \mu(\tau)\,p_L(\tau)\right) \nu ln(\gamma)\,.$$

In the following policy counterfactuals where θ,c_F, and η are changed from their calibrated values, an increase in θ is capable of producing very large and uniform increases in g for values of σ below the monopolization region. This occurs despite an endogenous response in the invariant distribution $\mu(\tau)$ that shifts weight away from the neck-and-neck state $\tau = 0$, where frontier innovation flows are the largest. Decreases in cost parameters c_F and η can increase growth close to the monopolization boundary through changes to the distribution $\mu(\tau)$. In particular, a decrease in c_F shrinks the monopolization region and introduces competitive innovation and growth.

3.6.1 The Effect of Increasing Innovator Bargaining Power θ

Increasing innovator bargaining power above a baseline of $\theta = 0$ has the effect of raising the gains from innovation, all else being equal. In particular, a firm at the frontier with technology lead $\tau \geq 0$ can sell their lead for more to the follower at $\tau + 1$, provided that $\theta > 0$.

From the first order condition for $p_L(\tau)$, I collect terms into a difference that expresses the above intuition of selling a technological lead for more at $\tau + 1$ versus τ:

$$\frac{\theta\left(\gamma^{\nu-1} f(\tau+1)\frac{S(\tau+1)}{r} - f(\tau)\frac{S(\tau)}{r}\right)}{c}.$$

For most values of σ below the monopolization boundary, surplus grows in τ, and $f(\tau)$ mirrors $S(\tau)$. Therefore the difference $\frac{\theta\left(\gamma^{\nu-1} f(\tau+1)\frac{S(\tau+1)}{r} - f(\tau)\frac{S(\tau)}{r}\right)}{c}$ is typically positive. An analogous intuition applies to the term that can be seen in the first order condition for p_E:

$$\frac{\theta\gamma^{\nu-1} f(1)\frac{S(1)}{r}}{c}.$$

The terms $2\left(\frac{\theta\gamma^{\nu-1}f(1)\frac{S(1)}{r}}{c}\right)$ and $\frac{\theta\left(\gamma^{\nu-1}f(\tau+1)\frac{S(\tau+1)}{r}-f(\tau)\frac{S(\tau)}{r}\right)}{c}$ perform reasonably well as predictors for the change in frontier innovation rates at $\tau = 0$ and $\tau > 0$, respectively, when θ is increased from zero. This occurs despite the fact that they do not fully reflect all dependencies present in the dynamic model. I perform a comparison of these predictors against the actual changes in frontier innovation rates for an increase in θ from 0 to 0.10, keeping the increase small to limit knock-on effects. This comparison is depicted in Figure 3.5. Although the predictor terms do not uniformly conform to actual changes across all σ and τ, they nevertheless are useful for illustrating the underlying mechanism behind the increases in innovation rates.

A discussion of the effect of θ on frontier innovation would not be complete without highlighting the impact on the follower's investment in imitation contact $f(\tau)$. The follower faces a holdup problem - if the innovator's bargaining power θ is high, then the follower is not incentivized to invest in imitation contact ex-ante. The expected payoff of the leader from knowledge transfer can be rewritten to incorporate the follower's first order condition for $f(\tau)$:

$$
\begin{aligned}
f(\tau)\theta S(\tau) &= \frac{(1-\theta)S(\tau)}{\eta}\theta S(\tau) \\
&= \frac{(1-\theta)\theta(S(\tau))^2}{\eta}.
\end{aligned}
$$

Holding $S(\tau)$ fixed, this expression shows that the maximum increase in frontier innovation rates will occur when θ is increased from 0 to $\frac{1}{2}$. Both innovator and imitator have a part to play in creating the market for knowledge; the innovator must be incentivized to create new technology, whereas the imitator must be incentivized to overcome the search frictions associated with knowledge transfer.

An increase in frontier innovation rates along with a decrease in $f(\tau)$ together cause the invariant distribution to shift weight onto higher τ, away from the neck-and-neck state $\tau = 0$ where frontier innovation rates are the highest because both firms are at the frontier. This endogenous response of the distribution $\mu(\tau)$ detracts from the resulting increase in g to a significant degree and even creates a net negative impact at values of σ close to the monopolization boundary. These effects are illustrated in Figure 3.7 for an increase from $\theta = 0$ to $\theta = 0.5$. In this figure, the dotted line represents an increase in θ to 0.5, holding fixed the original distribution. The dotted line conveys the strong impact of increased frontier innovation on growth, especially at low σ where surplus is high. However, the dashed line represents the full impact of both changes in innovation and the endogenous distribution shift. The resulting net effect on g is considerably smaller. On balance, however, an increase in θ from 0 to 0.5 is good for growth.

3.6.2 Decreasing Follower's Costs of Innovation and Imitation Contact

An incremental decrease in the follower's cost of innovation c_F will affect neck-and-neck innovation $p_E = \frac{\gamma v_L(1) - v_E}{c}$ and growth. In the Appendix Section 3.9.6 I use a simplified model to obtain the the following set of qualitative results, which occur together at sufficiently high substitutability. The first is that cost savings from c_F will create an upstream increase in value v_E at the neck-and-neck state. The second is that the flow rate of follower innovation $p_F(1)$ increases. The third is that the increase in $p_F(1)$ will cause a decrease in the value of the leader $v_L(1)$. Together, these imply that at higher degrees of substitutability, the gap $\gamma v_L(1) - v_E$ shrinks and p_E falls. Growth may decrease.

The first result, that v_E increases, also implies the possibility that a larger, discrete increase in c_F could shrink the monopolization region. This would occur because a large increase in v_E could change surplus $S(\tau)$ from being negative to positive at high τ. Followers that are far behind the frontier can now take advantage of knowledge transfer to catch up, and the monopolization region shrinks.

An incremental decrease in the follower's cost of imitation contact η will also affect p_E and growth. At low substitutability, cost savings from η will create an upstream increase in v_E and shrink the gap $\gamma v_L(1) - v_E$. The result is that p_E falls at low σ, and growth may decrease.

In Figure 3.8 I compare the change in growth resulting from large decreases in follower costs c_F and η against the effect of raising θ. The figure highlights the fact that changes in θ are capable of producing comparatively large and relatively uniform effects on g across σ.

This is true even for a increase of θ from 0 to 0.1. By contrast, reductions of c_F from 10.5 to 6 and η from 1.5 to 0.8 produce changes in growth of comparable or larger size only in the neighborhood of the monopolization boundary.

Aside from the informative comparison between policies, the figure largely confirms the intuition presented earlier about the relationship between follower costs, neck-and-neck innovation, and growth. In particular, reductions in c_F and η reduce growth at high and low substitutability, respectively. However, a reduction in c_F does shrink the monopolization region by raising v_E, as previously predicted. This results in a very large increase in growth at and above the former monopolization boundary. A reduction in η also creates growth just below the monopolization boundary because the invariant distribution is more concentrated on the neck-and-neck state where frontier innovation flows are highest.[11]

3.7 Conclusion

In this paper, I incorporate bargained knowledge transfer into the strategic innovation and diffusion framework exemplified by Aghion et al. (2001). I find that the model can produce a non-monotonic relationship of knowledge transfer to substitutability that is consistent with the empirical findings in Chapter 2. In particular, this correspondence between model and data provides suggestive evidence that the bargaining power of licensors is low in practice. In intellectual property policy counterfactuals, I find that the bargaining power is the most uniform and effectual instrument in raising growth, although it is limited by a countervailing shift in industry structure away from neck-and-neck innovation.

A couple of directions for future investigation emerge from these findings. The first concerns innovator bargaining power. Taken together, the empirical pattern and theoretical model suggest that bargaining power may be low in practice, and the dynamic model predicts a very large response in growth from a moderate increase in that parameter. This motivates further empirical and theoretical investigation into the determinants of bargaining power, whether in a technology licensing or an inventor-poaching environment.

[11] A much larger reduction in η is required to shrink the monopolization region.

The second direction concerns richer industry structure, whether in terms of more than two firms or incorporating entry and exit. However, the duopoly setup in this paper is able to model competition coming from other firms through its demand assumptions, and dovetails with the bilateral empirical measures that are employed in Chapter 2. The parsimonious approach enables a transparent and valuable intuition into the forces that shape the relationship of knowledge transfer to appropriability and competition.

3.8 Tables and Figures

Figure 3.1: Average Arrival Rate of Knowledge Transfer

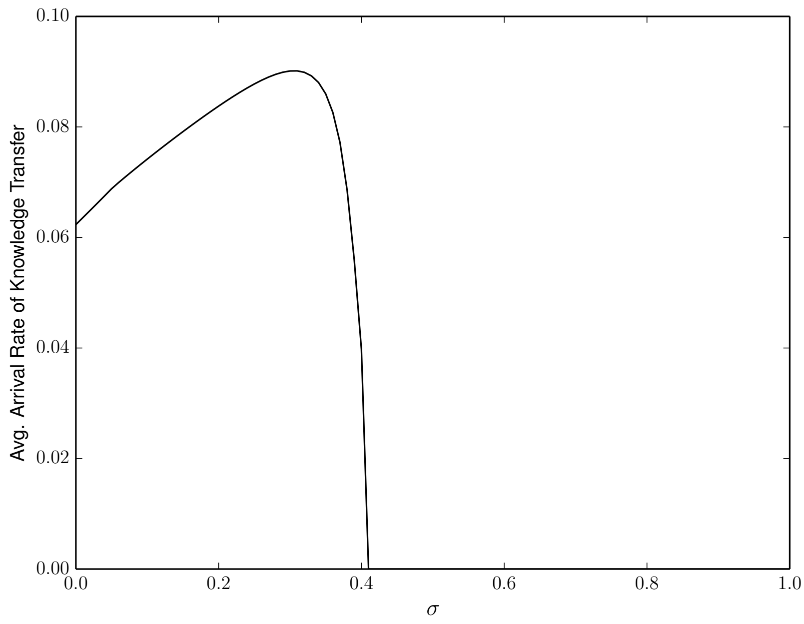

Figure 3.2: Decomposition of Knowledge Transfer

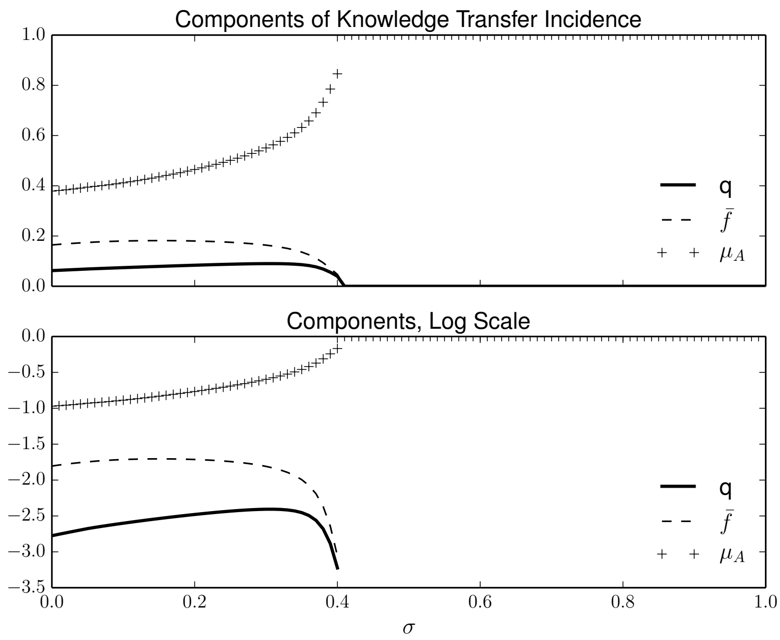

Figure 3.3: Decomposition of Opportunities for Knowledge Transfer

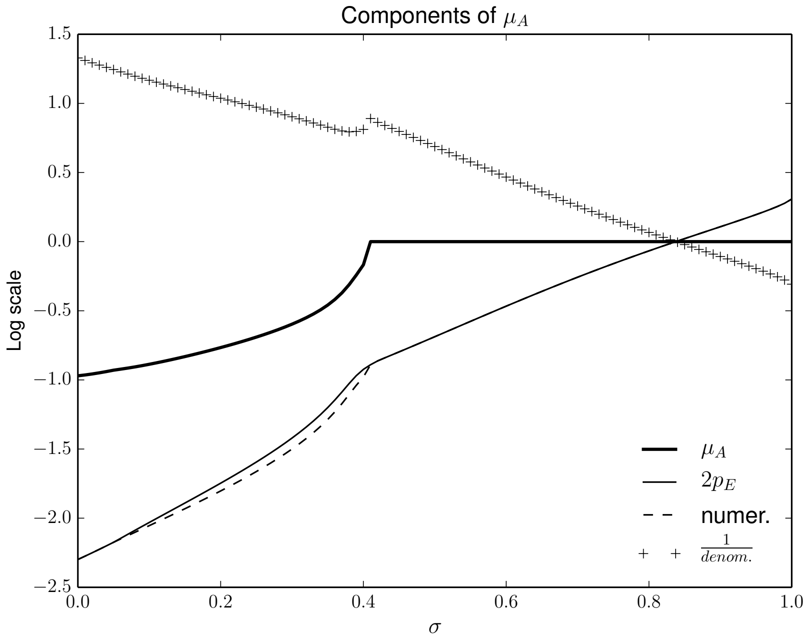

Figure 3.4: Leader vs. Follower Innovation at High Substitutability

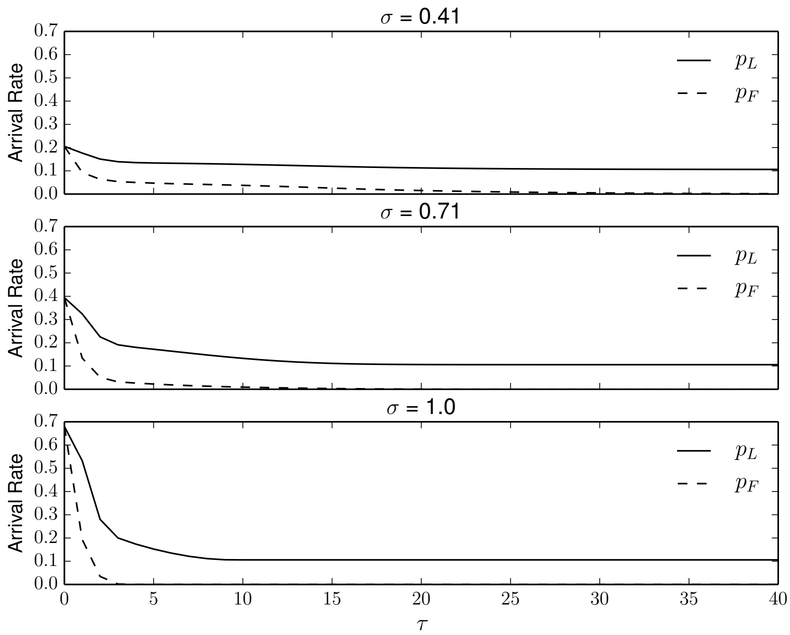

Figure 3.5: Actual vs. Predicted Changes in Frontier Innovation

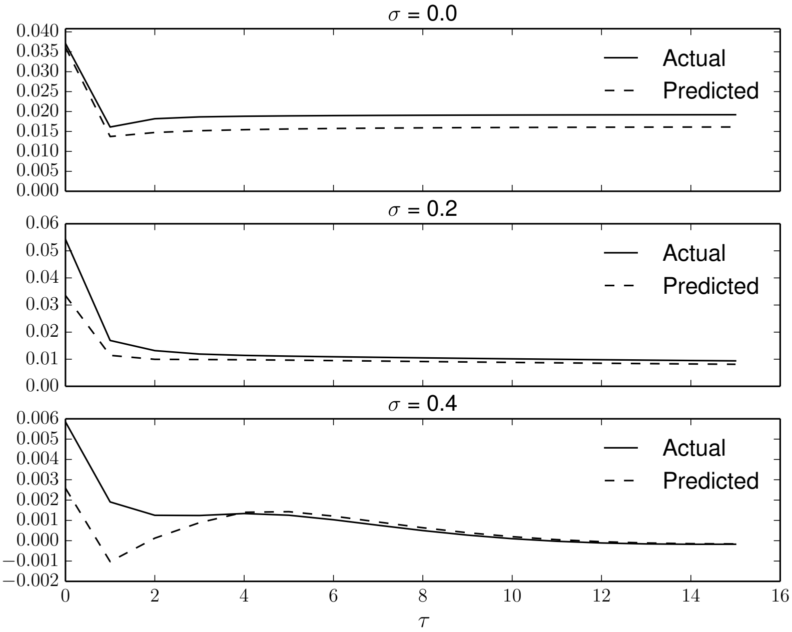

Figure 3.6: Knowledge Transfer for Different Values of Bargaining Power

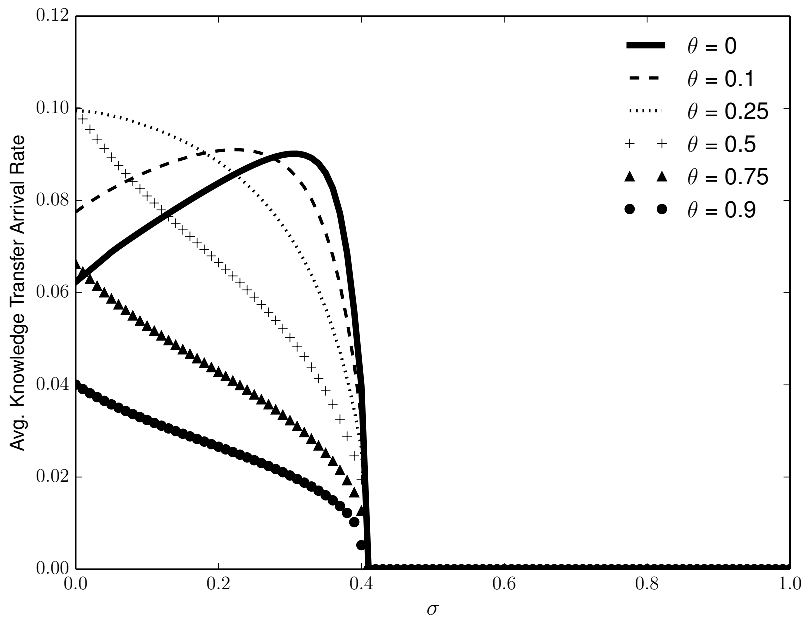

Figure 3.7: Growth, as Bargaining Power Increases

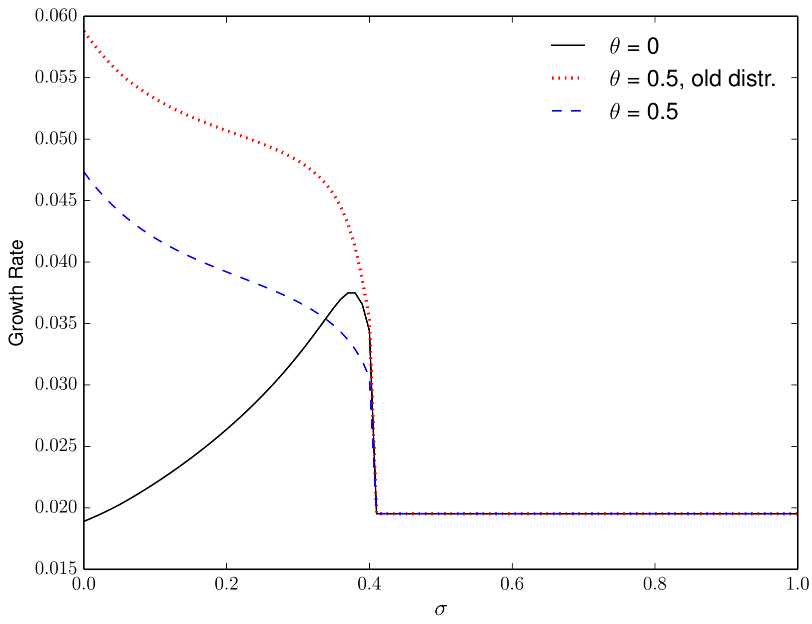

Figure 3.8: Growth, Across Three Types of Policy Changes

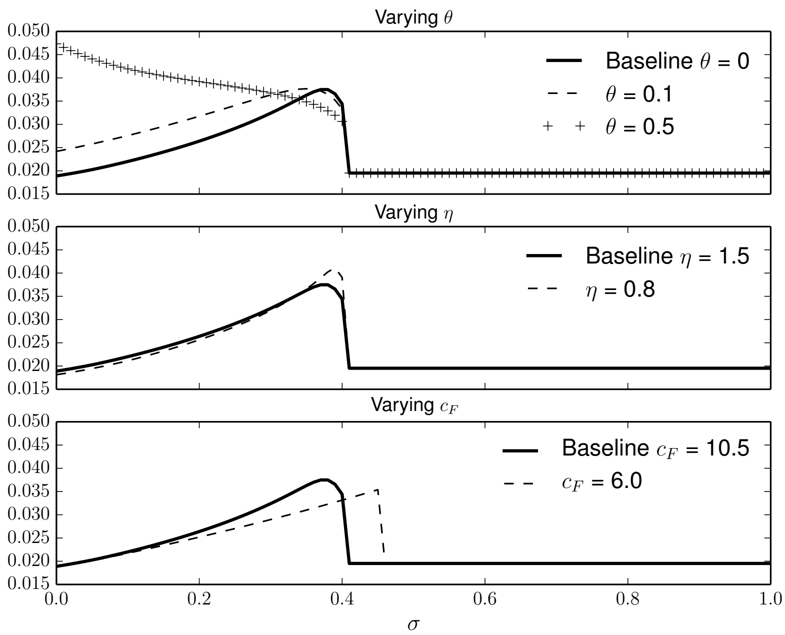

Table 3.1: Calibration Results

Parameter	Description	Calibrated Value	Target	Data Value	Model Value
γ	Growth factor	1.14	Avg. Growth Rate	2.6%	2.3%
η	Imitation contact cost	1.5	Variance of Growth Rate	0.25%	0.31%
c	Frontier innovation cost	6.0	Avg. IP invest/Nom. Output	3%	5.6%
c_F	Follower innovation cost	10.5	Var. IP invest/Nom. Output	0.02%	0.019%
$\frac{\nu-1}{\nu}$	Outer CES parameter	0.29	σ at which transfer peaks	0.25	0.31

3.9 Appendix

3.9.1 Appendix 1: Multiplicative Separability of Nested CES Indirect Profit

In Chapter 1, I show that static profit can be written as:

$$\pi_i = \frac{R_i}{\rho + \nu R_i} A \left(\frac{(\rho - 1) + (\nu - 1) R_i}{(\rho + \nu R_i)} \left(1 + \frac{1}{R_i}\right)^{\frac{1}{\rho - 1}} \right)^{\nu - 1} H_i^{\nu - 1}.$$

Recall that $R_i = \left(\frac{p_{-i}}{p_i}\right)^{\rho - 1}$ can be solved for in a fixed point problem as follows:

$$R_i = \left(\frac{\rho R_i + \nu}{(\rho - 1) R_i + (\nu - 1)} \frac{(\nu - 1) R_i + (\rho - 1)}{\nu R_i + \rho} \right)^{\rho - 1} \left(\frac{H_i}{H_{-i}} \right)^{\rho - 1}.$$

Suppose that firm i is the leader, such that $H_i = H$ and $\frac{H_i}{H_{-i}} = \gamma^{\tau}$. Then it is clear that R_L is a function of relative position τ only:

$$R_L = \left(\frac{\rho R_L + \nu}{(\rho - 1) R_L + (\nu - 1)} \frac{(\nu - 1) R_L + (\rho - 1)}{\nu R_L + \rho} \right)^{\rho - 1} (\gamma^{\tau})^{\rho - 1}.$$

The leader's profit $\pi(H_i, H_{-i})$ can be written as

$$\begin{aligned} \pi(H_i, H_{-i}) &= \frac{R_L}{\rho + \nu R_L} A \left(\frac{(\rho - 1) + (\nu - 1) R_L}{(\rho + \nu R_L)} \left(1 + \frac{1}{R_L}\right)^{\frac{1}{\rho - 1}} \right)^{\nu - 1} H^{\nu - 1} \\ &= \pi_L(\tau) H^{\nu - 1} \end{aligned}$$

where $\pi_L(\tau)$ is a function of τ only, through R_L:

$$\pi_L(\tau) = \frac{R_L}{\rho + \nu R_L} A \left(\frac{(\rho - 1) + (\nu - 1) R_L}{(\rho + \nu R_L)} \left(1 + \frac{1}{R_L}\right)^{\frac{1}{\rho - 1}} \right)^{\nu - 1}.$$

Conversely, if firm i is the follower, then R_F is also a function of τ only:

$$R_F = \left(\frac{\rho R_F + \nu}{(\rho-1)\,R_F + (\nu-1)} \frac{(\nu-1)\,R_F + (\rho-1)}{\nu R_F + \rho}\right)^{\rho-1} \left(\gamma^{-\tau}\right)^{\rho-1}.$$

The follower's profit $\pi_i\,(H_i, H_{-i})$ can be written as:

$$\begin{aligned}
\pi\,(H_i, H_{-i}) &= \frac{R_F}{\rho+\nu R_F} A \left(\frac{(\rho-1)+(\nu-1)\,R_F}{(\rho+\nu R_F)}\left(1+\frac{1}{R_F}\right)^{\frac{1}{\rho-1}}\right)^{\nu-1} \left(\gamma^{\nu-1}\right)^{-\tau} H^{\nu-1} \\
&= \pi_F\,(\tau)\,H^{\nu-1}
\end{aligned}$$

where $\pi_F\,(\tau)$ is a function of τ only, both directly as well as indirectly through R_F:

$$\pi_F\,(\tau) = \frac{R_F}{\rho+\nu R_F} A \left(\frac{(\rho-1)+(\nu-1)\,R_F}{(\rho+\nu R_F)}\left(1+\frac{1}{R_F}\right)^{\frac{1}{\rho-1}}\right)^{\nu-1} \left(\gamma^{\nu-1}\right)^{-\tau}.$$

Lastly, if both firms are equal, then $R_E = 1$ and profit is described as follows:

$$\begin{aligned}
\pi\,(H_i, H_{-i}) &= \pi_E H^{\nu-1}, \\
\pi_E &= \frac{A}{\rho+\nu}\left(\frac{(\rho-1)+(\nu-1)}{\rho+\nu} 2^{\frac{1}{\rho-1}}\right)^{\nu-1}.
\end{aligned}$$

Finally, here I provide solutions for the case where $\sigma = 1$ and correspondingly, $\rho \uparrow \infty$. A leader with H sets the aggregate price index P as the solution to the following problem:

$$\begin{aligned}
\max_P \quad & A P^{-\nu}\left(P - \frac{1}{H}\right) \\
s.t. \quad & \\
P \leq \; & \frac{\gamma^{\tau}}{H}.
\end{aligned}$$

Suppose the leader is constrained in their price-setting such that $P = \frac{\gamma^{\tau}}{H}$. Then profit is given by:

$$\pi(H_i, H_{-i}) = AH^{\nu-1}(\gamma^{\tau})^{-\nu}(\gamma^{\tau}-1).$$

Suppose on the other hand that they are not constrained because τ is sufficiently large. Then the first order condition gives:

$$P = \frac{1}{H}\frac{\nu}{\nu-1}.$$

That yields a corresponding monopoly profit of:

$$\pi(H_i, H_{-i}) = \frac{A}{\nu}H^{\nu-1}\left(\frac{\nu-1}{\nu}\right)^{\nu-1}.$$

3.9.2 Appendix 2: Multiplicative Separability of Value Functions

Using $V_L(\tau, H)$ as an example:

$$\begin{aligned}
rV_L(H,\tau) &= \max_{p_L(\tau)} \pi_L(\tau) H^{\nu-1} - H^{\nu-1}\frac{cp_L(\tau)^2}{2} \\
&\quad + p_L(\tau)\left(V_L(\gamma H, \tau+1) - V_L(H,\tau)\right) \\
&\quad + p_F(\tau)\left(\left(1_{\tau>1}V_L(H,\tau-1) + 1_{\tau=1}V_E(H)\right) - V_L(H,\tau)\right) \\
&\quad + f(\tau)\theta\bar{S}(H,\tau) \\
s.t. & \\
\bar{S}(H,\tau) &= \max\{2V_E(H) - V_L(H,\tau) - V_F(H,\tau), 0\}
\end{aligned}$$

I guess and verify that there exists a solution

$$v_i(\tau) = \frac{V_i(H,\tau)}{H^{\nu-1}}.$$

Substituting in the guess on both sides and factoring out $H^{\nu-1}$ on the right hand side:

$$\begin{aligned}
r v_L(\tau) H^{\nu-1} &= H^{\nu-1} \max_{p_L(\tau)} \{\pi_L(\tau) - \frac{c p_L(\tau)^2}{2} \\
&\quad + p_L(\tau)\left(\gamma^{\nu-1} v_L(\tau+1) - v_L(\tau)\right) \\
&\quad + p_F(\tau)\left((1_{\tau>1} v_L(\tau-1) + 1_{\tau=1} v_E) - v_L(\tau)\right) \\
&\quad + f(\tau)\theta S(\tau)\} \\
s.t. & \\
\bar{S}(\tau) &= \max\{2v_E - v_L(\tau) - v_F(\tau), 0\}.
\end{aligned}$$

Dividing through by $H^{\nu-1}$ on both sides, I show that $v_L(\tau)$ is recursively defined in only one state variable τ as follows:

$$\begin{aligned}
r v_L(\tau) &= \max_{p_L(\tau)} \{\pi_L(\tau) - \frac{c p_L(\tau)^2}{2} \\
&\quad + p_L(\tau)\left(\gamma^{\nu-1} v_L(\tau+1) - v_L(\tau)\right) \\
&\quad + p_F(\tau)\left((1_{\tau>1} v_L(\tau-1) + 1_{\tau=1} v_E) - v_L(\tau)\right) \\
&\quad + f(\tau)\theta S(\tau)\} \\
s.t. & \\
\bar{S}(\tau) &= \max\{2v_E - v_L(\tau) - v_F(\tau), 0\}.
\end{aligned}$$

Similar arguments apply for $V_F(\tau, H) = v_F(\tau) H^{\nu-1}$ and $V_E(H) = v_E H^{\nu-1}$.

3.9.3 Appendix 3: Asymptotic Static Profit and Value Function Behavior

Static Profit:

To arrive at $\lim_{\tau\uparrow\infty}\pi_L(\tau)$, first examine the asymptotic behavior of R_L:

Rearrange:

$$\frac{R_L}{\left(\gamma^{\rho-1}\right)^{\tau}}=\left(\frac{\rho R_L+\nu}{(\rho-1)R_L+(\nu-1)}\frac{(\nu-1)R_L+(\rho-1)}{\nu R_L+\rho}\right)^{\rho-1}$$

In Chapter 1, I prove that the right hand side, hereafter abbreviated RHS, has a negative slope everywhere in R_L, that $RHS|_{R_L=1}=1$, and therefore the left hand side (LHS) and RHS will intersect once at a solution of $R_L>1$.

As $\tau\uparrow\infty$, the R_L at which LHS and RHS intersect also goes to ∞, since $\lim_{R_L\uparrow\infty}RHS=\left(\frac{\rho(\nu-1)}{\nu(\rho-1)}\right)^{\rho-1}<1$, while the LHS rotates downward to have a slope in R_L that approaches zero .

Now, evaluating $\pi_L(\tau)$ as $\tau\uparrow\infty$ entails evaluating it as $R_L\uparrow\infty$:

$$\begin{aligned}\lim_{\tau\uparrow\infty}\pi_L(\tau) &= \lim_{R_L\uparrow\infty}\frac{R_L}{\rho+\nu R_L}A\left(\frac{(\rho-1)+(\nu-1)R_L}{\rho+\nu R_L}\left(1+\frac{1}{R_L}\right)^{\frac{1}{\rho-1}}\right)^{\nu-1}\\ &= \frac{A}{\nu}\left(\frac{\nu-1}{\nu}\right)^{\nu-1}\end{aligned}$$

To derive $\lim_{\tau\uparrow\infty}\pi_F(\tau)$, first examine R_F. By definition, $R_F=\frac{1}{R_L}$. So the following result obtains:

$$\lim_{\tau\uparrow\infty}R_F=0.$$

Returning to evaluate $\pi_F(\tau)$:

$$\begin{aligned}\lim_{\tau\uparrow\infty}\pi_F(\tau) &= \lim_{\tau\uparrow\infty}\left(\gamma^{\nu-1}\right)^{-\tau}\lim_{R_F\downarrow 0}\frac{R_F}{\rho+\nu R_F}A\left(\frac{(\rho-1)+(\nu-1)R_F}{\rho+\nu R_F}\left(1+\frac{1}{R_F}\right)^{\frac{1}{\rho-1}}\right)^{\nu-1}\\ &= 0\cdot A\left(\frac{\rho-1}{\rho}\right)^{\nu-1}\lim_{R_F\downarrow 0}\frac{(R_F+1)^{\frac{\nu-1}{\rho-1}}}{\rho+\nu R_F}\lim_{R_F\downarrow 0}R_F^{\frac{\rho-\nu}{\rho-1}}\\ &= 0.\end{aligned}$$

Now that I have established that $\pi_L(\tau)$ asymptotes to $\frac{A}{\nu}\left(\frac{\nu-1}{\nu}\right)^{\nu-1}$ and $\pi_F(\tau)$ to 0, I examine the asymptotic value function behavior.

Set $\tau=\bar{\tau}$ for $\bar{\tau}$ sufficiently large such that $\pi_L(\bar{\tau})\approx\lim_{\tau\uparrow\infty}\pi_L(\tau)=\frac{A}{\nu}\left(\frac{\nu-1}{\nu}\right)^{\nu-1}$ and $\pi_F(\bar{\tau})\approx\lim_{\tau\uparrow\infty}\pi_F(\tau)=0$. At such a $\bar{\tau}$, it is approximately true that

$$\pi_L(\bar{\tau}-1)=\pi_L(\bar{\tau})=\pi_L(\bar{\tau}+1)=\frac{A}{\nu}\left(\frac{\nu-1}{\nu}\right)^{\nu-1}$$

and

$$\pi_F(\bar{\tau}-1)=\pi_F(\bar{\tau})=\pi_F(\bar{\tau}+1)=0.$$

I guess (and later verify in the value function iteration) that it is approximately true that

$$\begin{aligned}v_L(\bar{\tau}-1)&=v_L(\bar{\tau})&&=v_L(\bar{\tau}+1),\\ v_F(\bar{\tau}-1)&=v_F(\bar{\tau})&&=v_F(\bar{\tau}+1),\\ p_L(\bar{\tau}-1)&=p_L(\bar{\tau})&&=p_L(\bar{\tau}+1),\\ p_F(\bar{\tau}-1)&=p_F(\bar{\tau})&&=p_F(\bar{\tau}+1)=0,\\ f(\bar{\tau}-1)&=f(\bar{\tau})&&=f(\bar{\tau}+1).\end{aligned}$$

These guesses yield a leader value function at $\tau=\bar{\tau}$ of

$$
\begin{aligned}
rv_L(\bar{\tau}) &= \max_{p_L(\bar{\tau})} \frac{A}{\nu}\left(\frac{\nu-1}{\nu}\right)^{\nu-1} - \frac{cp_L(\bar{\tau})^2}{2} \\
&\quad + p_L(\bar{\tau})\left(\gamma^{\nu-1}-1\right)v_L(\bar{\tau}) \\
&\quad + f(\bar{\tau})\theta S(\bar{\tau}) \\
s.t. & \\
\bar{S} &= \max\{2v_E - v_L(\bar{\tau}) - v_F(\bar{\tau})\}.
\end{aligned}
$$

They yield a follower value function at $\tau = \bar{\tau}$ of

$$
\begin{aligned}
rv_F(\bar{\tau}) &= \max_{f(\bar{\tau})} - \frac{\eta f(\bar{\tau})^2}{2} \\
&\quad + p_L(\bar{\tau})\left(\gamma^{\nu-1}-1\right)v_F(\bar{\tau}) \\
&\quad + f(\bar{\tau})(1-\theta)S(\bar{\tau}) \\
s.t. & \\
\bar{S} &= \max\{2v_E - v_L(\bar{\tau}) - v_F(\bar{\tau})\}.
\end{aligned}
$$

3.9.4 Appendix 4: Intuition for Neck and Neck Innovation

In the two-state dynamic model, there are only two relative states: $\tau \in \{0, 1\}$. The only action that occurs in the asymmetric state $\tau = 1$ is an exogenous arrival rate $f(1)$ of bargained imitation. As before, the absolute state variable H can be removed and a solution exists as a function of τ only.

Value functions are given by:

$$
\begin{aligned}
rv_E &= \max_{p_i} \pi_E - \frac{cp_i^2}{2} + p_i\left(\gamma^{\nu-1} v_L(1) - v_E(1)\right) + p_{-i}\left(\gamma^{\nu-1} v_F(1) - v_E(1)\right),\\
rv_L(1) &= \pi_L(1) + f(1)\,\theta S(1),\\
rv_F(1) &= \pi_F(1) + f(1)(1-\theta) S(1).
\end{aligned}
$$

In equilibrium, innovation p_E is given by:

$$
p_E = \frac{\gamma^{\nu-1} v_L(1) - v_E}{c}.
$$

Rewrite v_E as a function of the equilibrium p_E, and expand $v_L(1)$ and $v_F(1)$:

$$
\begin{aligned}
rv_E &= \pi_E - \frac{cp_E^2}{2} + p_E\left(\gamma^{\nu-1}\left(\frac{\pi_F(1) + \pi_L(1) + f(1) S(1)}{r}\right) - 2v_E\right)\\
v_E &= \frac{\pi_E - \frac{cp_E^2}{2} + p_E \gamma^{\nu-1}\left(\frac{\pi_F(1)+\pi_L(1)+f(1)S(1)}{r}\right)}{r + 2p_E}
\end{aligned}
$$

Expanding v_L and v_E inside p_E:

$$
p_E = \frac{\gamma^{\nu-1}\frac{\pi_L(1)+f(1)\theta S(1)}{r} - \frac{\pi_E - \frac{cp_E^2}{2} + p_E\gamma^{\nu-1}\left(\frac{\pi_F(1)+\pi_L(1)+f(1)S(1)}{r}\right)}{r+2p_E}}{c}.
$$

The interest is in $\theta = 0$, so applying and rearranging:

$$
\begin{aligned}
p_E &= \frac{\gamma^{\nu-1}\frac{\pi_L(1)}{r}\left(1 - \frac{p_E}{r+2p_E}\right) - \frac{\pi_E - \frac{cp_E^2}{2} + p_E\gamma^{\nu-1}\left(\frac{\pi_F(1)+f(1)S(1)}{r}\right)}{r+2p_E}}{\alpha}\\
&= \frac{\frac{1}{r+2p_E}\left(\left(\gamma^{\nu-1}\pi_L(1) - \pi_E\right) + \gamma^{\nu-1} p_E\left(\frac{\pi_L(1)-\pi_F}{r} - \frac{f(1)S(1)}{r}\right) + \frac{cp_E^2}{2}\right)}{c}.
\end{aligned}
$$

Before exploring surplus and its relation to σ, I rewrite v_E in an alternate fashion:

$$
\begin{aligned}
rv_E &= \pi_E - \frac{cp_E^2}{2} + p_E\left\{\left\{\gamma^{\nu-1}v_L(1) + \gamma^{\nu-1}v_F(1) - 2\gamma^{\nu-1}v_E\right\} + 2\gamma^{\nu-1}v_E - 2v_E\right\} \\
&= \pi_E - \frac{cp_E^2}{2} - p_E\gamma^{\nu-1}S(1) + 2p_E v_E\left(\gamma^{\nu-1} - 1\right) \\
v_E &= \frac{\pi_E - \frac{cp_E^2}{2} - p_E\gamma^{\nu-1}S(1)}{r - 2p_E\left(\gamma^{\nu-1} - 1\right)}.
\end{aligned}
$$

I solve for surplus:

$$
\begin{aligned}
S(1) &= 2v_E - v_F - v_L \\
&= 2\frac{\pi_E - \frac{cp_E^2}{2} - p_E\gamma^{\nu-1}S(1)}{r - 2p_E\left(\gamma^{\nu-1} - 1\right)} - \frac{\pi_F(1) + \pi_L(1) + f(1)S(1)}{r} \\
&= \frac{1}{\left(1 + \frac{2p_E\gamma^{\nu-1}}{r - 2p_E\left(\gamma^{\nu-1}-1\right)} + \frac{f(1)}{r}\right)}\left(2\frac{\pi_E - \frac{cp_E^2}{2}}{r - 2p_E\left(\gamma^{\nu-1} - 1\right)} - \left(\frac{\pi_F(1) + \pi_L(1)}{r}\right)\right) \\
&= \frac{1}{\left(1 + \frac{2p_E\gamma^{\nu-1}}{r - 2p_E\left(\gamma^{\nu-1}-1\right)} + \frac{f(1)}{r}\right)} \\
&\quad \times\left(2\pi_E\left(\frac{1}{r - 2p_E\left(\gamma^{\nu-1} - 1\right)} - \frac{1}{r}\right) - \frac{cp_E^2}{r - 2p_E\left(\gamma^{\nu-1} - 1\right)} + \left(\frac{2\pi_E - \pi_F(1) - \pi_L(1)}{r}\right)\right).
\end{aligned}
$$

Taking the partial derivative with respect to σ, holding p_E constant:

$$
\frac{\partial S(1)}{\partial\sigma} = \frac{2\left(\frac{1}{r - 2p_E\left(\gamma^{\nu-1}-1\right)} - \frac{1}{r}\right)\frac{d\pi_E}{d\sigma} + \frac{1}{r}\frac{d(2\pi_E - \pi_F(1) - \pi_L(1))}{d\sigma}}{\left(1 + \frac{2p_E\gamma^{\nu-1}}{r - 2p_E\left(\gamma^{\nu-1}-1\right)} + \frac{f(1)}{r}\right)}.
$$

It's the case that static surplus has a negative slope $\frac{d(2\pi_E - \pi_F(1) - \pi_L(1))}{d\sigma} < 0$ and $\frac{d\pi_E}{d\sigma} < 0$, so the result obtains:

$$\frac{\partial S(1)}{\partial \sigma} < 0.$$

Having established $\frac{\partial S(1)}{\partial \sigma} < 0$, I can now address the sign of $\frac{\partial p_E}{\partial \sigma}$:

$$\frac{\partial p_E}{\partial \sigma} = \frac{1}{r + 2p_E} \frac{1}{c} \left(\frac{d}{d\sigma} \left(\gamma^{\nu-1} \pi_L (1) - \pi_E \right) + \frac{\gamma^{\nu-1} p_E}{r} \frac{d}{d\sigma} (\pi_L (1) - \pi_F (1)) - \frac{\gamma^{\nu-1} p_E f (1)}{r} \frac{\partial S (1)}{\partial \sigma} \right).$$

Because $\frac{\partial S}{\partial \sigma} < 0$, it is then the case that the third term $-\frac{\gamma^{\nu-1} p_E f(1)}{r} \frac{\partial S(1)}{\partial \sigma} > 0$.

The first term $\frac{d}{d\sigma} \left(\gamma^{\nu-1} \pi_L (1) - \pi_E \right)$ is equivalent to the term $\frac{d}{d\sigma} (\pi (1,0) - \pi (0,0))$ in Chapter 1. This was shown to be, to a first approximation, weakly positive.

Finally, the sign of the second term is determined by the sign of $\frac{d}{d\sigma} (\pi_L (1) - \pi_F (1))$, which I show using numerical solutions to be strictly positive in the wide parameter range $\frac{\nu-1}{\nu} \in [0.1, 0.5], \gamma \in [1.01, 1.40]$.

3.9.5 Appendix 5: Calibration of Model Targets

The dynamic model equivalents of the targets are first calculated for a given value of σ, then averaged across all $\sigma \in [0, 1]$.

Let $\frac{d}{dt} log (Q_t) = \frac{\frac{dQ}{dt}}{Q} \approx \frac{\Delta log(Q)}{\Delta t}$ be the percentage growth rate. When multiplied by the expected number of $\Delta log (Q)$ increments, the Δt in the denominator cancels. Instead of modeling cycles, I will simply count all events, because they all result in growth.

Let the growth events be classified into three types. There is forward movement (+) which moves the frontier forward. There is backward movement (−) which closes the technological gap. Within backward movement there is follower innovation, and there is follower imitation.

Let $P \hat{(} \tau)$ denote the aggregate price index at relative state τ, when the leader's H has been normalized to 1. The change in $log (Q_t)$ be represented thus:

$$
\begin{aligned}
\Delta log\left(Q\right) &= log\left(Q'\right)-log\left(Q\right) \\
&= log\left(\frac{AH'^{\nu}P\left(\hat{\tau'}\right)^{-\nu}}{AH^{\nu}P\left(\hat{\tau}\right)^{-\nu}}\right) \\
&= \nu\left(log\left(H'\right)-log\left(H\right)\right)-\nu\left(log\left(P\left(\hat{\tau'}\right)\right)-log\left(P\left(\hat{\tau}\right)\right)\right)
\end{aligned}
$$

Let $\Delta_+\left(\tau\right)$ be the forward change due to leader innovation from τ to $\tau' = \tau+1$. Note that there is an absolute level effect on growth, as well as a relative effect:

$$
\Delta_+\left(\tau\right) = \underbrace{\nu log\left(\gamma\right)}_{\text{Absolute}} \underbrace{-\nu\left(log\left(P\left(\hat{\tau+1}\right)\right)-log\left(P\left(\hat{\tau}\right)\right)\right)}_{\text{Relative}}.
$$

Let $\Delta_-\left(\tau\right)$ be the backward change due to follower innovation from τ to $\tau' = \tau-1$. Note that there is only a relative effect:

$$
\Delta_-\left(\tau\right) = \underbrace{-\nu\left(log\left(P\left(\hat{\tau-1}\right)\right)-log\left(P\left(\hat{\tau}\right)\right)\right)}_{\text{Relative}}.
$$

Note that a combination of leader innovation and subsequent follower innovation lead to cancellation of the relative effect:

$$
\Delta_+\left(\tau\right)+\Delta_-\left(\tau+1\right) = \nu log\left(\gamma\right).
$$

The above example illustrates the intuition for why average growth only depends on the absolute effect $\nu log\left(\gamma\right)$. If an invariant distribution exists, then the relative effect of every forward movement is cancelled by backward movement, whether through follower innovation or follower imitation. To see that this is true, let τ^+ indicate all states above τ. For an invariant distribution to exist, the forward flow $\mu\left(\tau\right)p_L\left(\tau\right)$ (the only source of inflow) from τ into this τ^+ metastate must be balanced by backward flow out of the τ^+ metastate.

Part of that backward flow (and cancellation) comes from immediate follower innovation from $\tau+1$: $\mu(\tau+1)\,p_F(\tau+1)$. But the rest of it is through the backward flow of imitation, given by $\sum_{\tau^+}\mu(\tau)\,f(\tau)$. The effect of imitation $\Delta_f(\tilde{\tau})$ that occurs at $\tilde{\tau}>\tau$ can be log-decomposed as follows:

$$\Delta_f(\tilde{\tau})=\sum_{s=1}^{s=\tilde{\tau}}\Delta_-(s)\,.$$

One of the terms in this imitation decomposition is the $\Delta_-(\tau+1)$ that cancels the relative portion of $\Delta_+(\tau)$.

After all these cancellations, what's left is simply the level growth effect $\nu log(\gamma)$ from forward movement. The calculation of average real industry output growth $E\left[\frac{d}{dt}log(Q_t)\right]$ can be approximated thus:

$$g\approx\left(\mu(0)\,2p_E+\sum_{\tau=1}^{\tau=\bar{\tau}}\mu(\tau)\,p_L(\tau)\right)\nu log(\gamma)\,.$$

This formula is analogous to that found in Aghion et al. (2001). The intuition is the same: average industry growth only depends on the innovation flows of firms at the frontier.

It turns out that this formula for g also holds when an invariant distribution does not exist, leading to monopolization and $\tau\uparrow\infty$. In such a scenario, only the leader firm is active, the computational model terminates the state space at a sufficiently large $\bar{\tau}$, the distribution is summarized by $\mu(\bar{\tau})=1$, and the leader invests in $p_L(\bar{\tau})>0$. In what follows, I demonstrate that $\Delta_+(\bar{\tau})\approx\nu log(\gamma)$; that is, the relative component converges to zero. This result enables the above formula for g to still hold in the presence of monopolization.

$$
\begin{aligned}
\lim_{\tau\uparrow\infty}\Delta_{+}(\tau) &= \lim_{\tau\uparrow\infty}\left(\nu log(\gamma)-\nu\left(log\left(P(\hat{\tau+1})\right)-log\left(P(\hat{\tau})\right)\right)\right) \\
&= \nu log(\gamma) \\
&\quad -\nu\lim_{\tau\uparrow\infty} log\left(\frac{\rho+\nu R_L(\tau+1)}{\rho-1+(\nu-1)R_L(\tau+1)}\right) \\
&\quad -\nu\lim_{\tau\uparrow\infty} log\left(\frac{\rho-1+(\nu-1)R_L(\tau)}{\rho+\nu R_L(\tau)}\right) \\
&\quad -\nu\lim_{\tau\uparrow\infty}\frac{1}{1-\rho} log\left(\frac{1+\frac{1}{R(\tau+1)}}{1+\frac{1}{R(\tau)}}\right). \\
&= \nu log(\gamma)-\nu\left(log\left(\frac{\nu}{\nu-1}\right)+log\left(\frac{\nu-1}{\nu}\right)+\frac{1}{1-\rho}log(1)\right) \\
&= \nu log(\gamma).
\end{aligned}
$$

Below I show that when when $\rho\uparrow\infty$, $\lim_{\tau\uparrow\infty}\Delta_{+}(\tau)$ is still equal to $\nu log(\gamma)$:

$$
\begin{aligned}
\lim_{\tau\uparrow\infty}\lim_{\rho\uparrow\infty}\Delta_{+}(\tau) &= \lim_{\tau\uparrow\infty} log\left(\frac{A\left(\frac{1}{\gamma H}\frac{\nu}{\nu-1}\right)^{-\nu}}{A\left(\frac{1}{H}\frac{\nu}{\nu-1}\right)^{-\nu}}\right) \\
&= \nu log(\gamma).
\end{aligned}
$$

The variance of output growth is more involved:

$$
\begin{aligned}
Var\left(\frac{d}{dt}ln(Q_t)\right) &= \mu(0)\,2p_E dt\left(\frac{\Delta_{+}(0)}{dt}\right)^2 \\
&\quad +\sum_{\tau>0}\mu(\tau)\,p_L(\tau)\,dt\left(\frac{\Delta_{+}(\tau)}{dt}\right)^2 \\
&\quad +\sum_{\tau>0}\mu(\tau)\,p_F(\tau)\,dt\left(\frac{\Delta_{-}(\tau)}{dt}\right)^2 \\
&\quad +\sum_{\tau>0}\mu(\tau)\,f(\tau)\,dt\left(\frac{\Delta_{f}(\tau)}{dt}\right)^2 \\
&\quad -g^2.
\end{aligned}
$$

In the calculation of the average and variance of the ratio of IP investment to nominal output, it should be emphasized that only expenditures from innovation costs are included. Imitation contact expenditures in the model are not classified as IP investment.

The nominal output can be written as follows:

$$PQ = AH^{\nu-1}\hat{P}^{1-\nu}.$$

Normalizing $A = 1$, the combined expenditure of leader and follower at state τ, as a ratio to nominal output, is as follows:

$$X(\tau) = \left(c\frac{p_L(\tau)^2}{2} + \frac{c_F}{(\gamma^{\nu-1})^{\tau}}\frac{p_F(\tau)^2}{2}\right)P\hat{(\tau)}^{\nu-1}.$$

3.9.6 Appendix 6: Intuition for Follower Cost Reductions

Here I use simplified two-state models to provide intuition for the impact of reductions in c_F and η.

First I analyze the effect of a reduction in c_F:

Define value functions as follows (I abuse notation by using γ to denote $\gamma^{\nu-1}$):

$$rv_E = \max_{p_i} \pi_E + p_i(\gamma v_L - v_E) + p_{-i}(\gamma v_F - v_E) - c\frac{p_i^2}{2}$$

$$rv_F = \max_{p_F} \pi_F + p_F(v_E - v_F) - \frac{c_F}{\gamma}\frac{p_F^2}{2}$$

$$rv_L = \pi_L + p_F(v_E - v_L)$$

Note that the static profit functions π_E, π_F, π_L are meant to capture aspects of value in the full dynamic model that have been abstracted from in this simplified context.

The change in v_E is given by:

$$\frac{dv_E}{dc_F} = \frac{p_E\left(\frac{dp_F}{dc_F}\frac{\gamma(v_E - v_L)}{r+p_F}\left(1 - \frac{v_E - \gamma v_F}{\gamma v_L - v_E}\right) - \frac{\frac{p_F^2}{2}}{r+p_F}\right)}{r + \left(1 - \frac{\gamma p_F}{r+p_F}\right)\left(\frac{2(\gamma v_L - v_E) + \gamma v_F - v_E}{c}\right)}.$$

The change in p_F is given by:

$$\frac{dp_F}{dc_F} = \frac{1}{c_F}\frac{1}{r+p_F}\left(\frac{dv_E}{dc_F}\gamma r - \left(rp_F + \frac{p_F^2}{2}\right)\right).$$

The change in v_L is given by:

$$\frac{dv_L}{dc_F} = \frac{1}{r+p_F}\left(p_F\frac{dv_E}{dc_F} + \frac{dp_F}{dc_F}(v_E - v_L)\right).$$

In the dynamic model, the v_E factor $\left(1 - \frac{v_E - \gamma v_F}{\gamma v_L - v_E}\right)$ decreases from 1 to roughly $\frac{1}{2}$ on an interval of σ lying below the monopolization boundary. Borrowing that characteristic here in the simple model, I arrive at the combination $\frac{dv_E}{dc_F} < 0, \frac{dp_F}{dc_F} < 0, \frac{dv_L}{dc_F} > 0$. Finally, those results leads to the following characterization of p_E at sufficiently high σ:

$$\frac{dp_E}{dc_F} = \frac{1}{c}\frac{d}{dc_F}(\gamma v_L - v_E) > 0.$$

Now, I analyze the effect of a reduction in η. Starting with value functions:

$$rv_E = \max_{p_i} \pi_E + p_i(\gamma v_L - v_E) + p_{-i}(\gamma v_F - v_E) - c\frac{p_i^2}{2}$$

$$rv_F = \max_{p_F} \pi_F - \eta\frac{f^2}{2} + fS$$

$$rv_L = \pi_L.$$

The impact on v_F arises from a cost reduction:

$$\frac{dv_F}{d\eta} = \frac{-\frac{f^2}{2}}{r - f\left(\frac{2p_E(\gamma-1)-r+\frac{(v_E-\gamma v_F)}{c}}{r+\frac{1}{c}(\gamma v_F - v_E)+2p_E}\right)}.$$

At low σ, the full dynamic model gives that $\gamma v_F - v_E \gg 0$. Borrowing that feature for the simple model gives the result that the denominator of this expression is positive and therefore $\frac{dv_F}{d\eta} < 0$.

This feeds through to v_E:

$$\frac{dv_E}{d\eta} = \frac{p_E \gamma \frac{dv_F}{d\eta}}{r + \frac{1}{c}(\gamma v_F - v_E) + 2p_E}.$$

At low σ, the above slope is also negative. Finally, the effect on p_E obtains:

$$\frac{dp_E}{d\eta} = -\frac{1}{c}\frac{dv_E}{d\eta} > 0.$$

Bibliography

Acemoglu, D. and U. Akcigit (2012). Intellectual property rights policy, competition and innovation. *Journal of the European Economic Association 10*(1), 1–42.

Aghion, P., N. Bloom, R. Blundell, R. Griffith, and P. Howitt (2005). Competition and innovation: An inverted-U relationship. *The Quarterly Journal of Economics 120*(2), 701–728.

Aghion, P., C. Harris, P. Howitt, and J. Vickers (2001). Competition, imitation and growth with step-by-step innovation. *The Review of Economic Studies 68*(3), 467–492.

Almeida, P. and B. Kogut (1999). Localization of knowledge and the mobility of engineers in regional networks. *Management Science 45*(7), 905–917.

Anand, B. N. and T. Khanna (2000). The structure of licensing contracts. *The Journal of Industrial Economics 48*(1), 103–135.

Arora, A. and A. Fosfuri (2003). Licensing the market for technology. *Journal of Economic Behavior & Organization 52*(2), 277–295.

Baslandze, S. (2015). Spinout entry, innovation, and growth. *Mimeo*.

Belenzon, S. and M. Schankerman (2013). Spreading the word: Geography, policy, and knowledge spillovers. *Review of Economics and Statistics 95*(3), 884–903.

Bessen, J. and E. Maskin (2009). Sequential innovation, patents, and imitation. *The RAND Journal of Economics 40*(4), 611–635.

Bloom, N., M. Schankerman, and J. Van Reenen (2013). Identifying technology spillovers and product market rivalry. *Econometrica 81*(4), 1347–1393.

Fallick, B., C. A. Fleischman, and J. B. Rebitzer (2006). Job-hopping in Silicon Valley: Some evidence concerning the microfoundations of a high-technology cluster. *The Review of Economics and Statistics 88*(3), 472–481.

Fierro, G., B. Balsmeie, G.-C. Li, K. Johnson, S. Pham, A. Kaulagi, B. Yeh, and L. Fleming (2014). Weekly disambiguations of US patent grants and applications. Technical report, Fung Technical Report No. 2014.05.12, College of Engineering, UC Berkeley.

Fosfuri, A. (2006). The licensing dilemma: Understanding the determinants of the rate of technology licensing. *Strategic Management Journal 27*(12), 1141–1158.

Fosfuri, A. and T. Rønde (2004). High-tech clusters, technology spillovers, and trade secret laws. *International Journal of Industrial Organization 22*(1), 45–65.

Garmaise, M. J. (2009). Ties that truly bind: Noncompetition agreements, executive compensation, and firm investment. *Journal of Law, Economics, and Organization 27*(2), 376–425.

Hamasaki, S. (2012). R&D competition with employee poaching. *Mimeo*.

Heggedal, T.-R., E. R. Moen, and E. Preugschat (2014). Productivity spillovers through labor mobility. *Centre for Economic Policy Research Discussion Paper No. DP9850*.

Irwin, D. A. and P. J. Klenow (1994). Learning-by-doing spillovers in the semiconductor industry. *Journal of Political Economy 102*(6), 1200–1227.

Kaiser, U., H. C. Kongsted, and T. Rønde (2015). Does the mobility of R&D labor increase innovation? *Journal of Economic Behavior & Organization 110*, 91–105.

Katz, M. L. and C. Shapiro (1985). On the licensing of innovations. *The RAND Journal of Economics 16*(4), 504–520.

Katz, M. L. and C. Shapiro (1987). R&D rivalry with licensing or imitation. *The American Economic Review 77*(3), 402–420.

Kim, Y. and N. S. Vonortas (2006a). Determinants of technology licensing: The case of licensors. *Managerial and Decision Economics 27*(4), 235–249.

Kim, Y. and N. S. Vonortas (2006b). Technology licensing partners. *Journal of Economics and Business 58*(4), 273–289.

Marx, M., D. Strumsky, and L. Fleming (2009). Mobility, skills, and the Michigan non-compete experiment. *Management Science 55*(6), 875–889.

Møen, J. (2005). Is mobility of technical personnel a source of R&D spillovers? *Journal of Labor Economics 23*(1), 81–114.

Pakes, A. and P. McGuire (1994). Computing Markov-perfect Nash equilibria: Numerical implications of a dynamic differentiated product model. *The RAND Journal of Economics 25*(4), 555–589.

Pakes, A. and S. Nitzan (1983). Optimum contracts for research personnel, research employment, and the establishment of "rival" enterprises. *Journal of Labor Economics 1*(4), 345–365.

Radauer, A. and T. Dudenbostel (2013). Patlice survey: Survey on patent licensing activities by patenting firms. *Directorate-General for Research and Innovation, EUR 26114*.

Shapiro, C. (1985). Patent licensing and R&D rivalry. *The American Economic Review 75*(2), 25–30.

Stoyanov, A. and N. Zubanov (2012). Productivity spillovers across firms through worker mobility. *American Economic Journal: Applied Economics 4*(2), 168–198.